DTIC FILE COPY

(4)

AD

Technical Memorandum 21-89

COMPARISON OF OCULOMETER AND HEAD-FIXED RETICLE WITH VOICE OR SWITCH AND TOUCH PANEL FOR DATA ENTRY ON A GENERIC TACTICAL AIR COMBAT DISPLAY

AD-A217 231

Christopher C. Smyth
Mary E. Dominessy

November 1989
AMCMS Code 612716.H700011

DTIC
ELECTE
JAN 30 1990
S B D

Approved for public release;
distribution is unlimited.

U.S. ARMY HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland 21005-5001

90 01 30 003

®DEC, PDP, UNIBUS, VAX, VMS are registered trademarks of Digital Equipment Corporation.

®Carroll Touch is a registered trademark of Carroll Touch Incorporated.

®Interstate VRM is a registered trademark of Interstate Electronics Corporation.

®NAC Eyemark Recorder is a registered trademark of NAC Incorporated.

®Polhemus Navigation Sciences Division 3Space and Polhemus Isotrak are registered trademarks of Polhemus Navigation - Sciences Division, McDonnell Douglas Electronics Company, McDonnell Douglas Corporation.

®SAS Statistical Analysis Computer Package is a registered trademark of SAS Institute Incorporated.

Destroy this report when no longer needed.
Do not return it to the originator.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Use of trade names in this report does not constitute an official endorsement or approval of the use of such commercial products.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution is unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
Technical Memorandum 21-89

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
Human Engineering Laboratory

6b. OFFICE SYMBOL *(If applicable)*
SLCHE

7a. NAME OF MONITORING ORGANIZATION

6c. ADDRESS *(City, State, and ZIP Code)*
Aberdeen Proving Ground, MD 21005-5001

7b. ADDRESS *(City, State, and ZIP Code)*

8a. NAME OF FUNDING / SPONSORING ORGANIZATION

8b. OFFICE SYMBOL *(If applicable)*

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS *(City, State, and ZIP Code)*

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
6.27.16	1L162716AH70		

11. TITLE *(Include Security Classification)*
Comparison of Oculometer and Head-Fixed Reticle With Voice or Switch and Touch Panel for Data Entry on A Generic Tactical Air Combat Display

12. PERSONAL AUTHOR(S)
Smyth, Christopher C., Dominessy, Mary E.

13a. TYPE OF REPORT
Final

13b. TIME COVERED
FROM ________ TO ______

14. DATE OF REPORT *(Year, Month, Day)*
1989, November

15. PAGE COUNT
77

16. SUPPLEMENTARY NOTATION
Keywords: (man computer interface)

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
01	03	01
23	02	

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Army air-to-air helicopters; Eyegaze pointing tasks; tactical air combat display; head-fixed reticle; Oculometers; touch panel

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

An experiment with 15 U.S. Army enlisted military subjects was conducted to compare the use of a head-mounted oculometer, a head-fixed reticle, and a touch panel for data entry tasks on a generic tactical air combat display. The oculometer and the fixed reticle were operated in a head-free mode and were used with either switch or voice. The fixed reticle with switch, the oculometer with switch, and the touch panel are significantly faster than the fixed reticle with voice and the oculometer with voice. The fixed reticle methods are faster than the oculometer methods of the same modality. The switch methods are faster than the voice methods. The ocular pointing methods (whether oculometer or fixed reticle, voice or switch input) require a larger display activation window (±1.1 inches at 27 inches' viewing distance) than does the touch panel, thereby limiting the number of selections that can be shown on the display. This is especially true for the oculometer with voice method which generated significantly more selection errors and may therefore require a still larger display activation window for proper operation.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
Technical Reports Office

22b. TELEPHONE *(Include Area Code)*
(301) 278-4478

22c. OFFICE SYMBOL
SLCHE-SS-IR

AMCMS Code 612716.H700011 Technical Memorandum 21-89

COMPARISON OF OCULOMETER AND HEAD-FIXED RETICLE WITH VOICE OR SWITCH AND TOUCH PANEL FOR DATA ENTRY ON A GENERIC TACTICAL AIR COMBAT DISPLAY

Christopher C. Smyth
Mary E. Dominessy

August 1989

APPROVED: John D. Weisz
JOHN D. WEISZ
Director
Human Engineering Laboratory

Approved for public release;
distribution is unlimited.

U.S. ARMY HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland 21005-5001

ACKNOWLEDGMENTS

Mr. Joseph F. Dunn, Service Support Division Test Branch (SSDTB), Human Engineering Laboratory (HEL), in consultation with the senior investigator, designed and built in the best tradition of the skilled craftsman, a head support system for the NAC Eyemark oculometer that enabled the successful conduct of this experiment.

DTIC
COPY
INSPECTED
1

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification

By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

CONTENTS

EXECUTIVE SUMMARY 3

INTRODUCTION 5

OBJECTIVE 7

METHODOLOGY 7

Apparatus 7
Test Subjects 28
Experimental Design 28
Training and Test Procedures 30

RESULTS 32

DISCUSSION 42

CONCLUSIONS 67

RECOMMENDATIONS FOR FURTHER RESEARCH 67

REFERENCES 69

APPENDIX

Posttest Questionnaire and Results 71

FIGURES

1. Eyemark Oculometer: Front View 10
2. Eyemark Oculometer: Right-Front View 11
3. Test Station 13
4. Subject Using Touch Panel 15
5. Subject Using Oculometer With Push Button 16
6. Tactical Display Screen 17
7. Display Screens for Data Request Task Using Switches as Input Modality 19
8. Display Screens for Identification Report Task Using Switches as Input Modality 20
9. Display Screens for Data Request Task Using Voice as Input Modality 21
10. Display Screens for Identification Report Task Using Voice as Input Modality 22
11. Display Screens for Data Request Task Using Touch Panel as Input Modality 23
12. Display Screens for Identification Report Task Using Touch Panel as Input Modality 24
13. Calibration of Oculometer 33
14. Calibration Pattern for Oculometer 34
15. Calibration Check of Oculometer 35

16. Task by Method Interaction for Track Select Task Time (T1) 47
17. Task by Method Interaction for Submenu Select Task Time (T2) 48
18. Task by Method Interaction for Subtask Completion Time (T3) 49
19. Task by Method Interaction for Total Task Time (TT) 50
20. Spatial Tolerance Frequency Distribution by Method 65

TABLES

1. Display Screen Size and Characteristics 18
2. Scenario Track Probabilities 26
3. Test Subject Demographic Data 29
4. Testing Sequence Assignments 36
5. Results of the Repeated Measures Multivariant Analysis of Variance for the Transformed Time Variables: Wilks Criterial and Mauchly's Criteria for Sphericity, the Correlation of the Means amd Variances, and the Normality Test 38
6. Univariate Analysis of Variances for Within-Subject Effects of the Transformed Time Variables 39
7. Scheffē's Test for the Methods 40
8. Chi-Squared Analysis of the Total Errors by Task, Method, and Task by Method Interaction (0.05 Significance Level) 41
9. Descriptive Statistics for the Task Times and Total Errors for the Method by Task Interaction: Number, Average, Standard Deviation, and Standard Error 43
10. Statistical Analysis of the Subjects' Preference Ratings by the Friedman Two-Way Analysis of Variance for Related Samples 45
11. Statistical Analysis of Subject Errors by Error Type 51
12. Statistical Analysis of the Selection Errors as a Function of Subtask 54
13. Statistical Analysis of the Relationship Between the Times and Errors of the Select-Track Subtask 57
14. Statistical Analyses of the Difference Between the Average Times for Error-Free Runs and Total Runs 60
15. Statistical Analysis of the Error-Free Times 61
16. Descriptive Statistics for the Spatial Tolerance of the Track Selection Subtask Number, Average, Minimum (MIN), Maximum (MAX), and Standard Deviation (SD) 63
17. Frequency Distribution of the Spatial Tolerance of Each Method 64
18. Descriptive Statistics of the Spatial Tolerance Distributions for the Methods 66

EXECUTIVE SUMMARY

An experiment with 15 U.S. Army enlisted military subjects was conducted to compare the use of a head-mounted oculometer, a head-fixed reticle, and a touch panel for data entry tasks on a generic tactical air combat display. The oculometer and the fixed reticle were used with either switch or voice. The experiment is a repeated measures (15x5x2) design with tasks and methods within subjects fixed factors, and subjects as a random variable.

The subjects were tested on the five device configurations in a counterbalanced scheme. They used the devices to perform a data entry task and a data extraction task in a randomly occurring manner, both tasks requiring interaction with the tactical display. The subjects selected the track symbol and menu items needed to complete the task specified by an instruction line shown on the display at the start of each test trial. The times and errors committed by the subject in completing the task were used to evaluate the five data entry methods.

The statistical analysis, a multivariate analysis of variance (MANOVA), shows a significant difference in performance for the five configurations. The fixed reticle and switch, the oculometer and switch, and the touch panel are significantly faster than the fixed reticle and voice and the oculometer and voice. The fixed reticle methods are faster than the oculometer methods of the same modality. The switch methods are faster than the voice methods. The ocular pointing methods (whether oculometer or fixed reticle, voice or switch input) require a larger display activation window (±1.1 inches at 27 inches' viewing distance) than does the touch panel, therefore limiting the number of selections that can be shown on the display. This is especially true for the oculometer and voice method that generated significantly more selection errors and may therefore require a still larger display activation window for proper operation.

COMPARISON OF OCULOMETER AND HEAD-FIXED RETICLE WITH VOICE OR SWITCH AND TOUCH PANEL FOR DATA ENTRY ON A GENERIC TACTICAL AIR COMBAT DISPLAY

INTRODUCTION

It is expected that high work loads will be imposed on pilots in single crew military helicopters during the air-to-air combat role. This is especially true during the target acquisition phase of air combat when the pilot is interacting with an on-board panel-mounted tactical air combat display used for target alerting and cuing. The pilot will need to specify the display tracks of interest to amplify flight data as an aid in selecting potential targets. The pilot may orient himself toward the target from the display. He must then acquire the target with the fire control sensor and designate fire engagement. He may have to visually confirm an enemy target before engagement depending on the rules of engagement. This process must take place as quickly and effectively as possible if the pilot is to survive over the battlefield.

Several alternate methods of display interaction are investigated as possible means of reducing this work load through reduced data entry times during display interaction. The methods are relatively novel soldier-display interface mechanisms based on the direct coupling of the visual process and the display control mechanism. One such method is based on the use of a head-fixed reticle for alignment with a display item of interest, and another on an oculometer to designate the eye gaze direction.

The head-fixed reticle concept has been incorporated into on-board helicopter fire control systems to designate visually acquired targets outside the aircraft for fire engagement. For example, the pilot of the Apache attack helicopter uses the Honeywell Integrated Helmet and Display Sighting System (IHADSS) to designate targets. The pilot moves his head so that the fixed reticle on the monocular helmet-mounted display is aligned with the image of the intended target. The sensor of the armament system slues with the head movements of the pilot, and the target is acquired by pushing a switch, or possibly (in the future) by speaking an engage fire command.

An alternate target designation method would employ oculometers. The Army and NASA have proposals under contract to develop head-mounted oculometers for target designation in the single pilot Light Helicopter Experimental (LHX) cockpit. In this case, the pilot needs only to look at the intended target as the oculometer measures his gaze direction and to press a switch or voice the appropriate command to acquire the target. The concept of using oculometers for military fire control is well established (Setterholm, 1982). A helmet-mounted eye tracker was invented for the Navy (Breglia, 1981) for fire control. The Air Force has supported research in using target tracking that employs eye movements with a Honeywell remote panel-mounted oculometer (Meyer, 1981). A head position sensor must be integrated with a helmet-mounted oculometer for extension to weapon system pointing and cockpit panel display control.

The on-board tactical air combat display, envisioned for the air-to-air helicopter as an early warning and cuing device, will be included in the Forward Area Air Defense (FAAD) automated system planned for the 1990s. The Army is currently working toward automating the transmission of track alerting and cuing and command and control (C^2) information to the aviation elements and air defense fire units. The information will appear on tactical displays within the FAAD command, control, and intelligence (C^2I) system at the battalion, battery, platoon, and fire unit level. Similar information would be sent to displays within the aviation command and on the flight elements.

The displays would show the aircraft detected by the FAAD system within the immediate area as symbols on a tactical map display. The displays would allow rapid communication of tactical information between units and headquarters. The high performance of modern aircraft necessitates the early alerting and cuing of the fire units and aviation elements about the presence of aircraft within their area.

An on-board air combat tactical display will benefit the Army air-to-air combat helicopter. The recent development of Army air-to-air combat helicopter capabilities is in response to the Soviet arming of the HIND helicopter as a defense of their tank forces against allied helicopter antiarmor attack. They have extended this effort to the development of air-to-air combat helicopter airframes. The Army air-to-air combat helicopter, either a redesigned Apache or the new LHX, would be used to defend allied ground attack helicopters. Another role, recently considered, is active air defense units in which air-to-air helicopters would be directed to intercept attacking enemy helicopters or aircraft before they reach the defending ground forces.

The inclusion of a tactical display on-board the air-to-air combat helicopter would allow the pilot the same advantage of early alert and cuing that the FAAD system affords the fire units. The pilot would be able to plan approach routes and select interception points from the display. He would be forewarned of the presence of enemy threats in the area and be able to take evasive action. Furthermore, the pilot will be able to interact with the display and report tactical intelligence to higher headquarters about enemy air activities of interest to both aviation and air defense elements.

The method that the helicopter pilot uses to retrieve information from the tactical display will partly determine its usefulness. This is certainly true for the single pilot LHX. The pilot must direct most of his attention to the primary task of continually flying the aircraft; the secondary task of display interaction should require little attention during air-to-air combat. This is especially true for helicopter displays with their small size. All information available about aircraft in the immediate area is stored in a track data file within an on-board computing processor. The information is updated every few seconds from the open digital data FAAD broadcast and the helicopter's satellite navigation system. However, all information about every track cannot be displayed simultaneously on the tactical display because of the limited size of the display. The pilot chooses the track symbol of interest from the display and then retrieves the track data information (wing type, airspeed, altitude, heading, number) needed to compute a reasonable intersection point. Similar comments apply to transmitting an intelligence message of the status of observed aircraft. The pilot interacts with the display by "capturing" ("hooking" in air defense nomenclature) the display

symbol of interest and then selecting the appropriate menu action. Capturing or hooking is defined as selecting a target symbol from the display and tagging it for further processing.

An effective method of display interaction would be similar in modality to that used by the pilot for target engagement to prevent confusion between tasks and reduce work load. A research question of interest is whether the pilot can use these ocular-based methods, developed for outside target designation, to interact with the on-board displays, particularly with the tactical air combat display.

OBJECTIVE

The purpose of this study is to compare the performance, as measured by the data entry times and errors, of an oculometer with switch or voice, a head-fixed reticle with switch or voice, and a touch panel as data entry methods on a generic tactical air combat display. The touch panel is included for comparison of the ocular-based methods to a standard display interactive device.

METHODOLOGY

The apparati, test subjects, experimental design, and training and test procedures are described in this section.

Apparatus

The apparati used in this experiment are listed below. The operating characteristics and limitations of the equipment that are pertinent to understanding the experimental procedures and results are described in this section. The following apparati were used in this experiment:

1. DEC VAX 11/780 host computer consisting of a central processing unit, a floating point accelerator, and 2.75 megabytes of memory. The computer has UNIBUS adapter interfaces to an Aydin display processing unit, real time clock, discrete digital input and output switches, analog-to-digital converters, and communications ports to VT100/220 terminals. The VAX Virtual Memory System (VMS) operating system supports the FORTRAN language in real time simulation of military systems. The VMS language provides the priority, scheduling, process creation and control, real time event-driven response, and high speed, interprocess communications essential for real time simulation of complicated systems.

2. Aydin graphics system (model 5216 display computer) providing a 1024- by 1024-pixel resolution. The system has five memory planes that can generate 16 simultaneous colors with an overlay for alphanumerics. The memory bus controller and processor controls vector and character generation, and permits pixel loading from the host computer at 800 pixels per second. The

refresh memory modules provide a 1024- by 1024- by 5-pixel storage resolution for the video output. Interface to the host computer is by a parallel DR11-W direct memory access UNIBUS.

3. Aydin model 8026 color graphics, video monitor driven by the Aydin raster scan graphics system (model 5216 display computer) providing a 1024- by 1024-pixel resolution. The monitor is a 19-inch diagonal (15.5 by 11 inches), high resolution, red-green-blue (RGB) color monitor for use with the Aydin graphics system.

4. Carroll Touch input system with an infrared scanning sensor for the Aydin 8026 video monitor. The 19-inch diagonal sensor provides a 64- by 48-opto-matrix of addressable points. Communication to the host computer is by an RS232C interface.

5. Interstate Voice Recognition Modular (VRM) automatic speech recognition system, a speaker-dependent, isolated speech, voice data entry peripheral with noise-canceling microphone for voice command entry. The recognizer communicates to the host computer by an RS232C interface. The user must supply a list of vocabulary words that will be spoken. The user enrolls these words on the recognizer to establish reference templates.

6. A custom-made keypad with "HOOK" and "CANCEL" keys. The keys are interfaced to the host computer through the DRS11/DSS11 input system. The input modules are the DSS11 series allowing contact sense input to be sent to the host computer. All DSS11 input signals are optically isolated for user protection. The keys are wired to the DSS11 interrupt line. A key-push will cause a service request interrupt to occur, forcing the host computer to poll the switches.

7. An NAC Eyemark recorder, Model V, with field camera unit (V-19), right and left eye mark camera units (V-15 and 16), a controller (V-71), a data output unit (V-99), and connecting cables.

8. A Polhemus Navigation Sciences Division 3Space Isotrak low frequency magnetic field position and orientation indicator with magnetic field source, sensor, and controller device.

9. A detachable plastic visor with sighting cross hairs for a head-fixed reticle.

10. A head support fixture for holding the Eyemark to the subject's head. The fixture acts as a frame for the Isotrak sensor, noise-canceling microphone, and detachable head-fixed reticle.

11. A large screen (25-inch diagonal) television monitor for viewing the Eyemark returns during the calibration and testing.

12. A wooden stand fixture with sighting device for locating the Aydin display in the 3Space coordinate system. The sighting device holds the Isotrak sensor during a calibration procedure.

13. A large wooden housing frame and a display console. Both structures were custom-built by the Human Engineering Laboratory (HEL). The wooden frame provides a fixed alignment between the display console and the Polhemus field source. The console holds the switch keys.

14. A VT100 computer terminal used to control the test program.

Described in the remainder of this section of the report are (a) the functioning of the Eyemark recorder and the Polhemus, (b) the equipment integration, (c) the integration of the signals from both devices, (d) the interfacing of the devices to the host computer, (e) the computer processes including the display driver, (f) the display concept, and (g) the size of the display data selection window as determined by the spatial accuracy of the devices.

a. Eyemark Recorder. The NAC Eyemark recorder is a head-mounted point of gaze recorder that employs three sub-miniature television cameras to record the field of view and the left and right eye focal points. A near infrared light source is used to illuminate the cornea of each eye with invisible light. Many oculometer designs are based on a measurement of infrared light reflected from the human eye; the NAC oculometer measures the corneal reflection. This is in contrast to the bright pupil technique used with the Honeywell remote oculometer and the SRD Ltd. head-mounted eye tracker, and the measurement of both the bright pupil and corneal reflection used with the Applied Science Laboratories' helmet-mounted oculometer to compensate for the effects of small shifts of the helmet on the head.

The reflected virtual image from each eye is transmitted by a mirror system in the camera units to a charge-coupled array and from there to a processor that automatically tracks the corneal reflection from each eye. The three TV signals are integrated along with the eye tracks and sent to the controller and a video display monitor. The coordinates of each of the eye tracks are forwarded to the data output unit for output to analog-to-digital converters. The coordinates are updated at a 30-Hertz (Hz) rate. This rate is high enough to monitor visual smooth pursuit, saccadic, and other visual search criteria; a monitoring rate of 3 to 6 Hz is sufficient for eye gaze control purposes, however.

b. Polhemus 3Space Isotrak. The Polhemus 3Space Isotrak low frequency magnetic field device determines the position and orientation of a sensor relative to a magnetic field source, thereby providing a full 6-degrees-of-freedom measuring device. The device allows continual monitoring of the position and orientation of the system to which the sensor is attached. The information is transmitted to a host computer by an RS232 interface in ASCII or binary format at a 9600-baud rate. The Isotrak configuration comprises a source, a sensor, and an electronics processing unit. The small magnetic sensor contains orthogonally wound coils in a 1/2-inch cube; the field source also contains orthogonally wound coils in a 1-inch cube. The source generates a low frequency (10.24 kHz) magnetic field that is measured by the sensor. The processor unit computes the position and orientation of the sensor relative to the source, and controls the transmission of the output data.

c. Equipment Integration. Figure 1 shows a front view and Figure 2 shows an oblique view of the Eyemark recorder mounted on a mannequin's head.

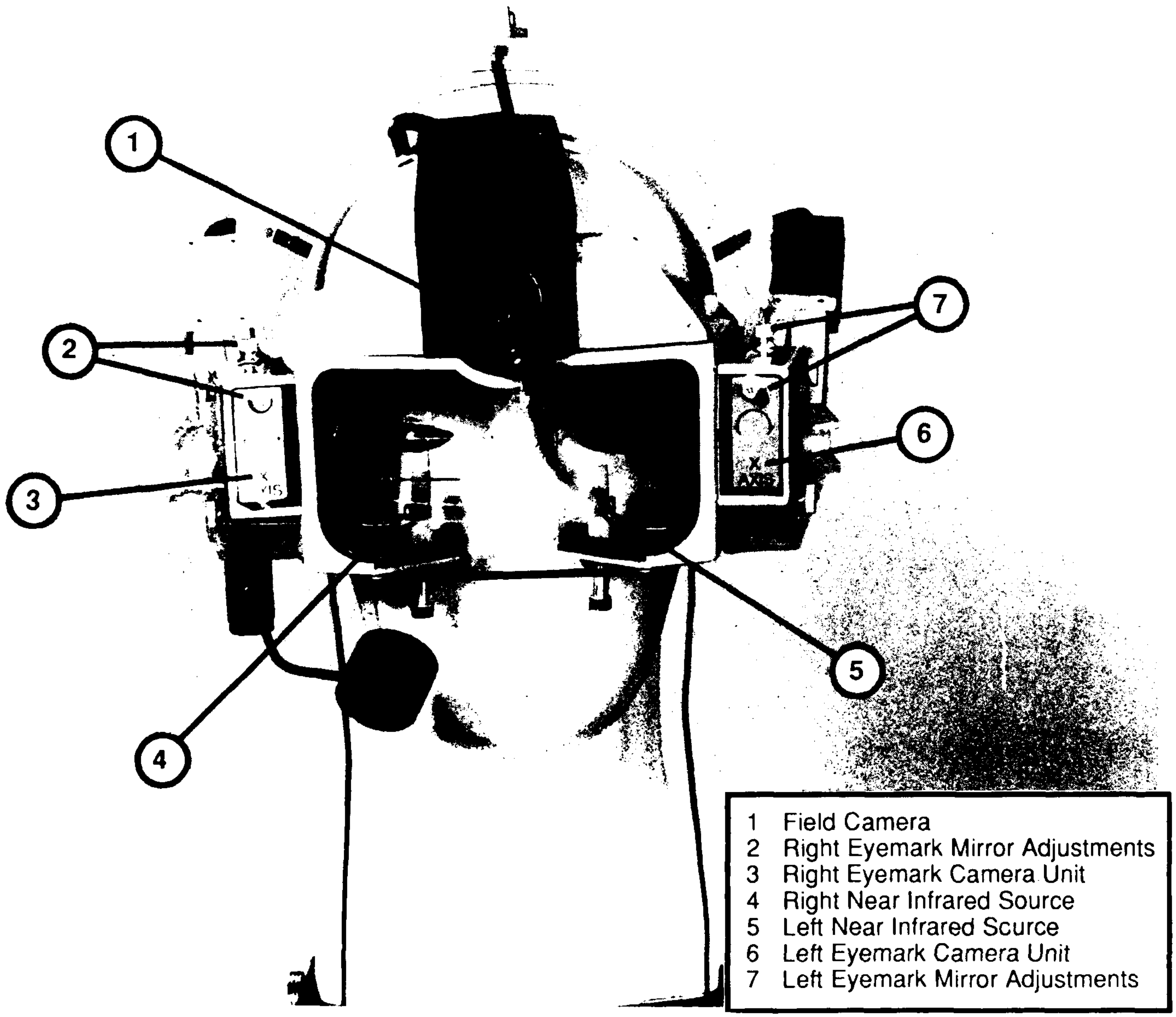


Figure 1. Eyemark oculometer: Front view.

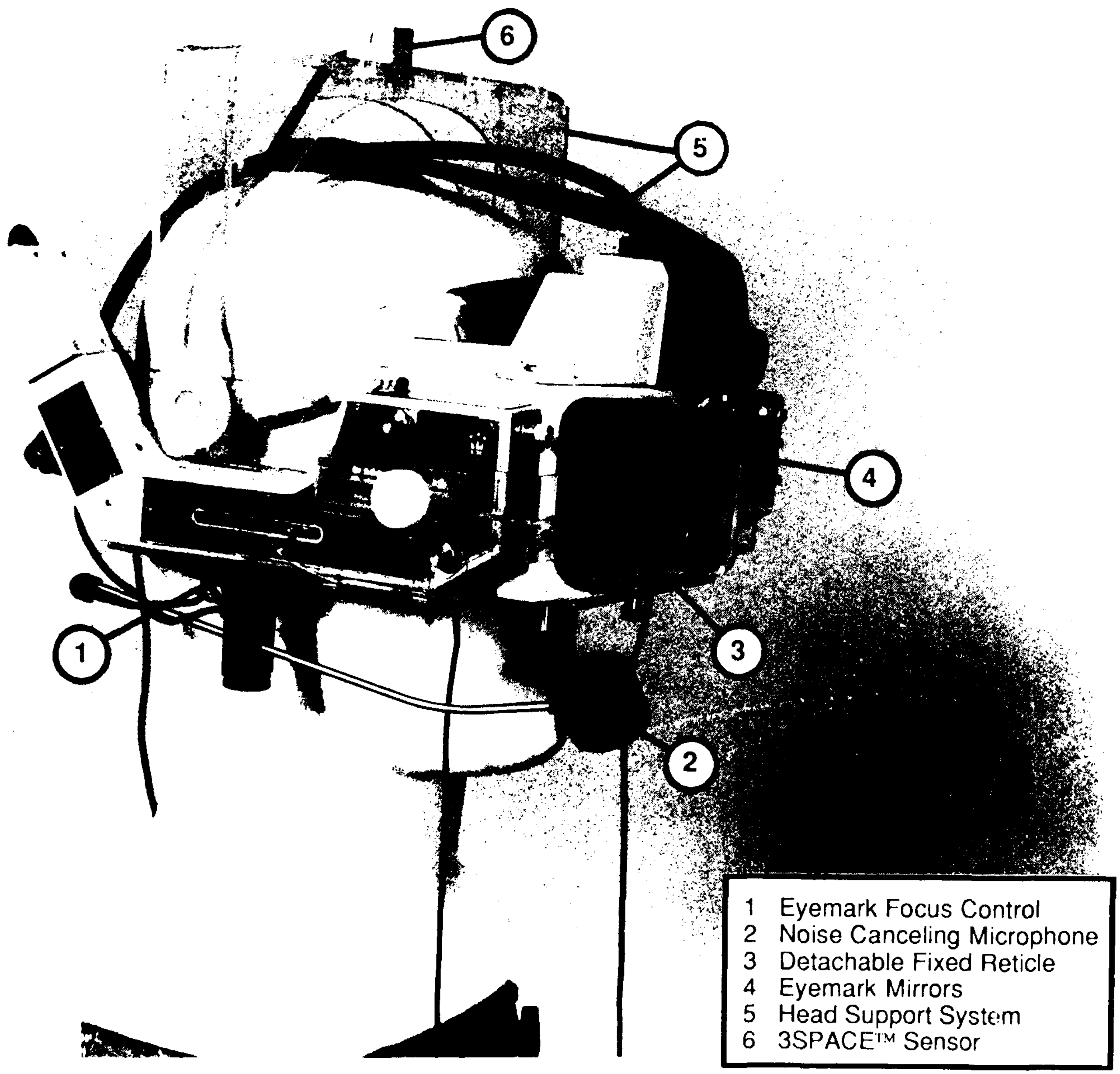


Figure 2. Eyemark oculometer: Right-front view.

The figures show the head support system for the recorder, the 3Space sensor mounted on top the support, and the detachable head-fixed reticle and noise-canceling microphone mounted on the right side of the support system.

Figure 3 shows a view of the experimental console and test apparatus. The experimental console contains the Aydin raster scan display screen mounted above an operator console panel with the keypad attached. The tactical display shown on the raster scan display is positioned between desk and eye level. The Carroll Touch infrared device is aligned with the surface of the display. The display console shelf is at desk height and contains the control panel with the keypads. The console was designed in accordance with MIL-STD-1472C (Department of Defense, 1981) for human engineering design criteria. Figure 3 shows the VT100 and Polhemus controller box to the left of the test console. The Eyemark controller box and data output unit are on the wooden stand to the right of the test console. The Interstate speech recognizer is on top the data output box. The TV monitor is to the right of the equipment stand. A large wooden frame attached to the test console provides a fixed alignment between the display console and the Polhemus field source, which is shown in the upper right of the figure.

The position and orientation of the Aydin display surface is located in the 3Space source coordinate system in a preliminary calibration process. This is done using a sighting tube mounted on a stand. The 3Space sensor is mounted on the sighting tube and returns the position and orientation of the sighting tube in the field source coordinate system. The experimenter sights along the tube at cue marks on the display face. The location and orientation of the display face is computed using a least squares regression analysis technique from the known locations of the cue marks in the 2-dimensional coordinate system of the display, the straight line distances from the cue marks to the sensor, and the 3Space sensor returns for each of the sightings.

d. Signal Integration. The subject can use the fixed reticle mounted on the head support fixture in conjunction with the 3Space sensor to designate items on the display. When the head is rotated so that the image of the displayed item is aligned with the cross hair reticle, the item may be selected by means of switch or voice. The position and orientation readings from the 3Space sensor are used to compute the subject's eye position and gaze direction in the 3Space coordinate system. The point where the gaze intersects the display is then computed from the known location and orientation of the display surface (see paragraph c, Equipment Integration). Since both the reticle and sensor are fixed to the head support system being worn by the subject, the position of the subject's eye and the viewing direction through the reticle's cross hairs are fixed relative to the sensor. The relation between these parameters and the sensor's position and orientation readings is established in a calibration process (see Training and Test Procedures section) before usage.

Furthermore, the subject can use the oculometer without the fixed reticle, in conjunction with the 3Space sensor, to designate items on the display simply by gazing at them. The oculometer must first be centered and aligned to the subject's vision field. The output from the oculometer is combined with the sensor output to compute an eye position and viewing direction for the subject in the 3Space coordinate system. The point where the gaze intersects the display is then computed from the known location and orientation of the display surface. Since both the oculometer and sensor are

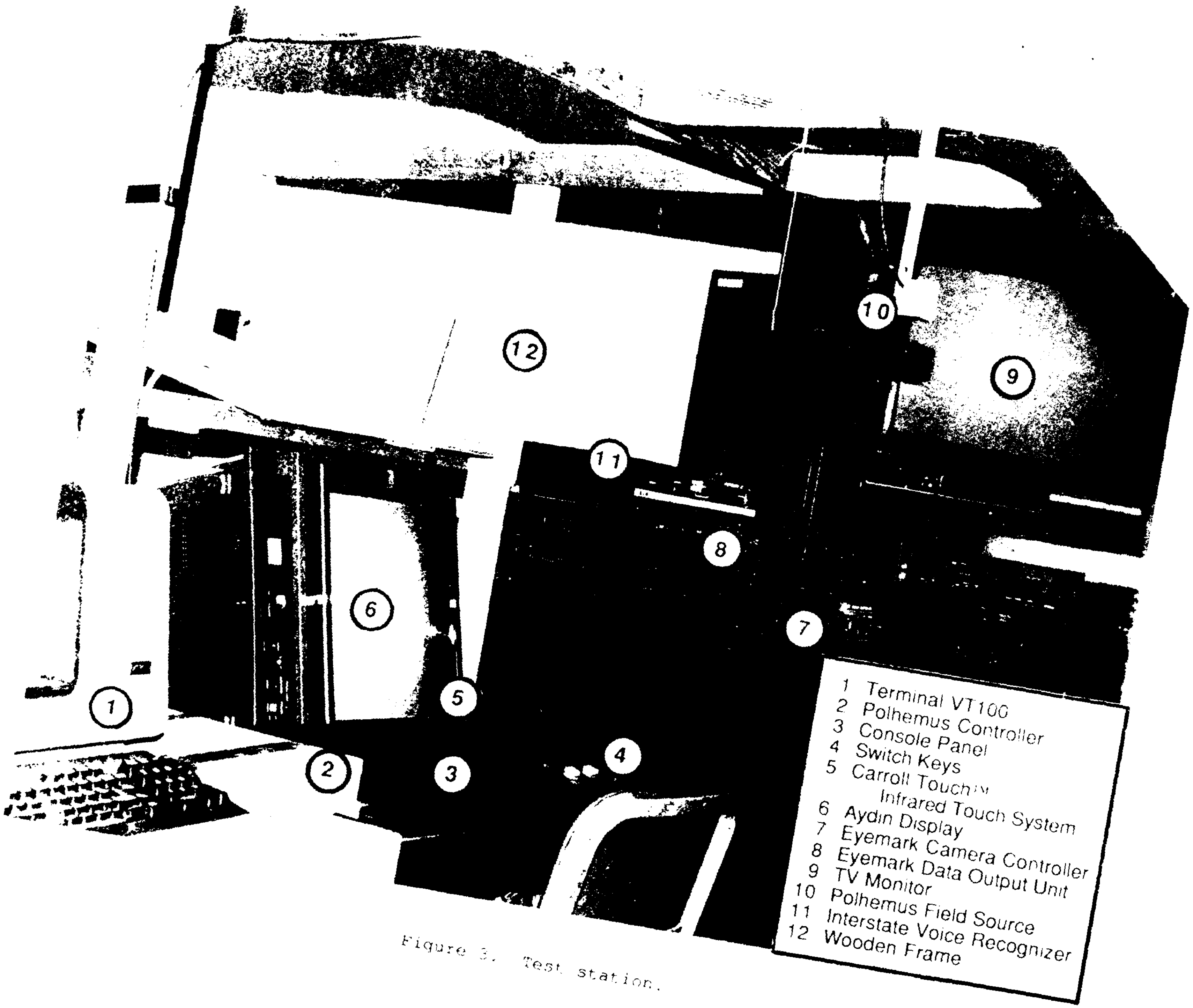


Figure 3. Test station.

fixed to the head support system being worn by the subject, the position of the subject's eye and the coordinate axes of the oculometer are fixed relative to the sensor. The relation between these parameters and the sensor's position and orientation readings is established in two calibration processes (see Training and Test Procedures) before usage.

e. Computer Interface. The test was conducted at the experimental console shown in Figure 3, and controlled by the experimenter from the DEC VT100 terminal placed to the side. The subject sits in front of the console and operates a tactical display shown on the Aydin monitor under computer program control driven by the DEC VAX 11/780. The subject interacts with the display using each of the experimental methods.

Figure 4 shows the subject using the Carroll Touch system. The subjects were instructed to use a wooden pencil to select display items via the touch panel. Experience has shown that subjects using a finger for data selection will inadvertently relax their hands enough to allow more than one finger to break the infrared beams of the touch panel therefore generating panel detection errors.

Figure 5 shows the subject wearing the Eyemark head support system. The figure appropriately depicts either the fixed reticle or the oculometer configuration when switches are used for data entry. Alternately, the subject could be speaking a command into the head-mounted microphone.

f. Process Control. The DEC VAX 11/780 computer is the process controller for all test phases. All equipment is interfaced to the computer. The Interstate VRM, Carroll Touch system, and Polhemus 3Space are interfaced via RS232 ports. The keypad is interfaced through the discrete digital input and output switches. The analog outputs for the "x" and "y" positions of the right and left eyes from the NAC Eyemark Recorder are sampled by the analog-to-digital converters at a 3-Hz rate. Separate computer program processes service each device; the routines communicate through a global common area.

A separate process drives the Aydin monitor showing the tactical display in Figure 6. The test scenario track symbols are updated once every second on the tactical display. The voice or switch entries, causing changes in the display in a selected track symbol or a menu, are serviced immediately for user feedback. Entries not corresponding to a selected item cause an error message to be displayed. Voice entries that are not recognized as a reference template, cause the VT100 to be momentarily "beeped" as feedback to the subject that a misrecognition has occurred.

The use of the VMS operating system is necessary for real time programming of complicated configurations if the response times of the different devices are to be reduced to match the sensitivity of the human subject. The VMS operating system allows the execution of different subprocesses servicing the different devices. The processes run independently under system level control, but exchange data through event flags and global common areas. The subprocesses can be controlled by various system level services to schedule the processing of events.

g. Display Concept. The characteristics of the tactical display are presented in Table 1. The display is divided into the instruction area, the graphics display area, and the menu display area. Figures 7 through 12

Figure 4. Subject using touch panel.

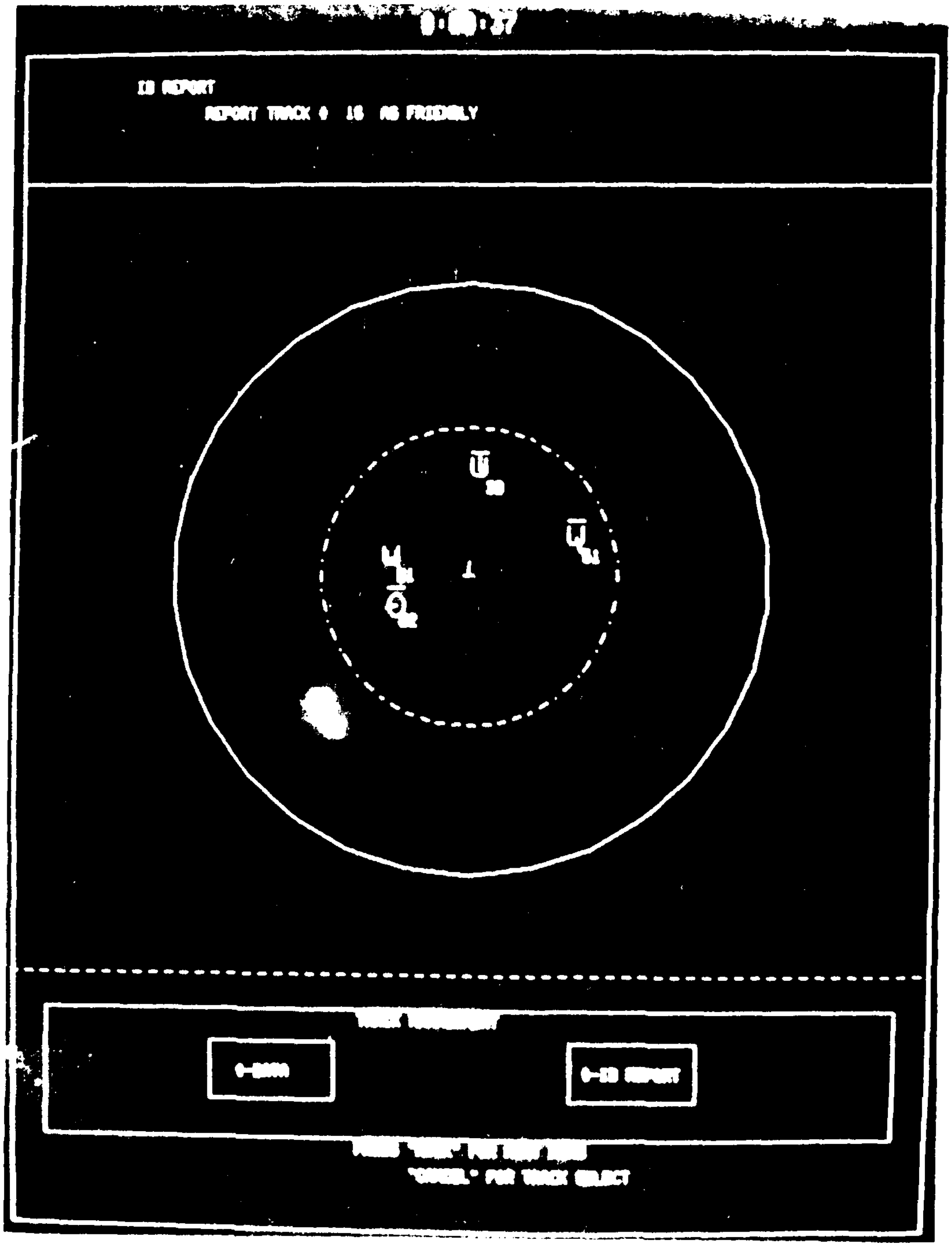


Figure 6. Tactical display screen.

Table 1
Display Screen Size and Characteristics

Display element	Characteristic	Value (inches)
Aircraft symbol	Size	0.228
Host aircraft	Size	0.142
Range ring	Radius	3.000
Alphanumeric	Size	
Track numbers		0.100
Menu characters		0.130
Screen	Size	
Width		10.000
Length		
Overall		13.000
Instruction area		1.500
Graphics area		8.500
Menu/Submenu area		3.000

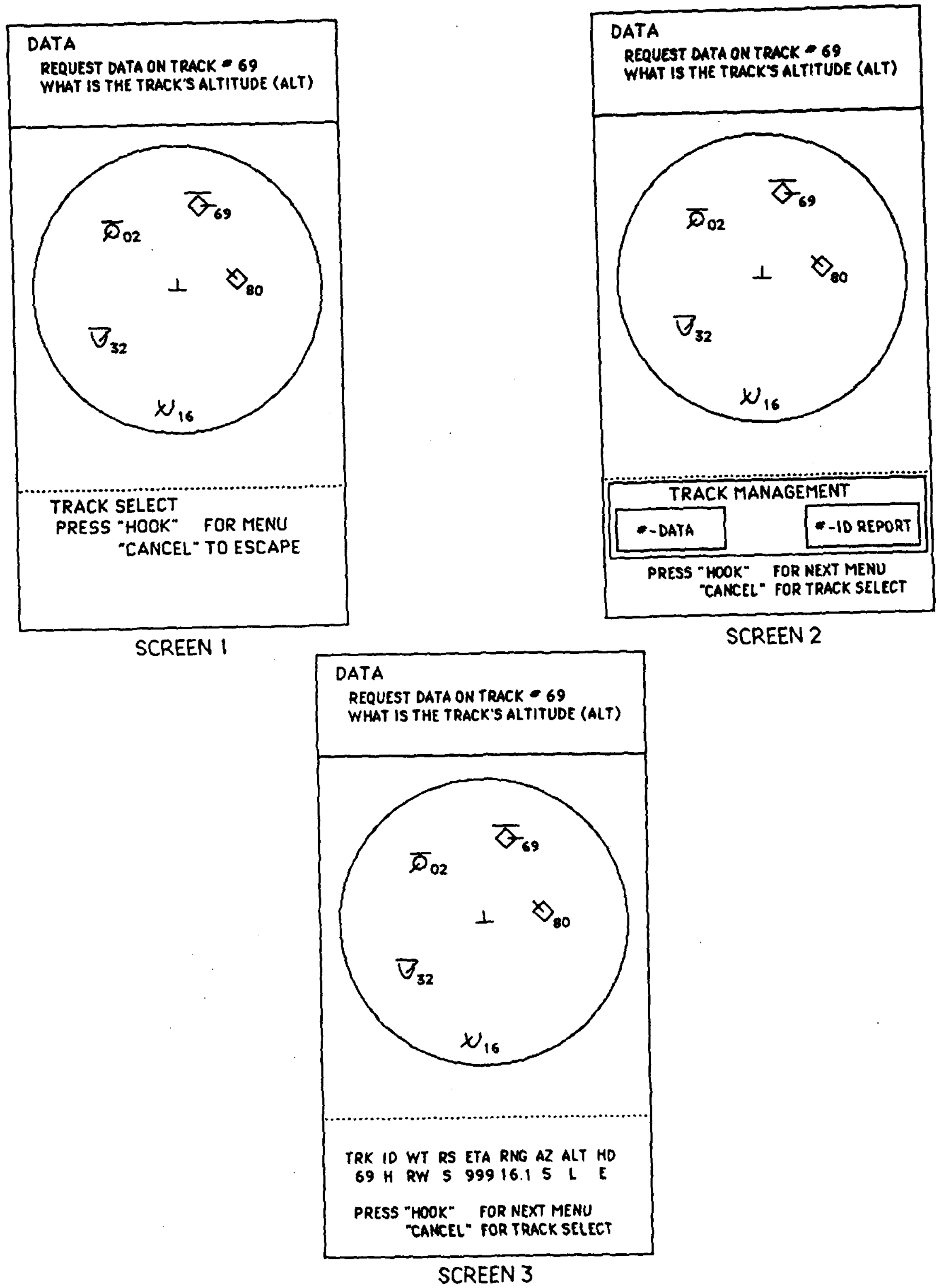


Figure 7. Display screens for data request task using switches as input modality.

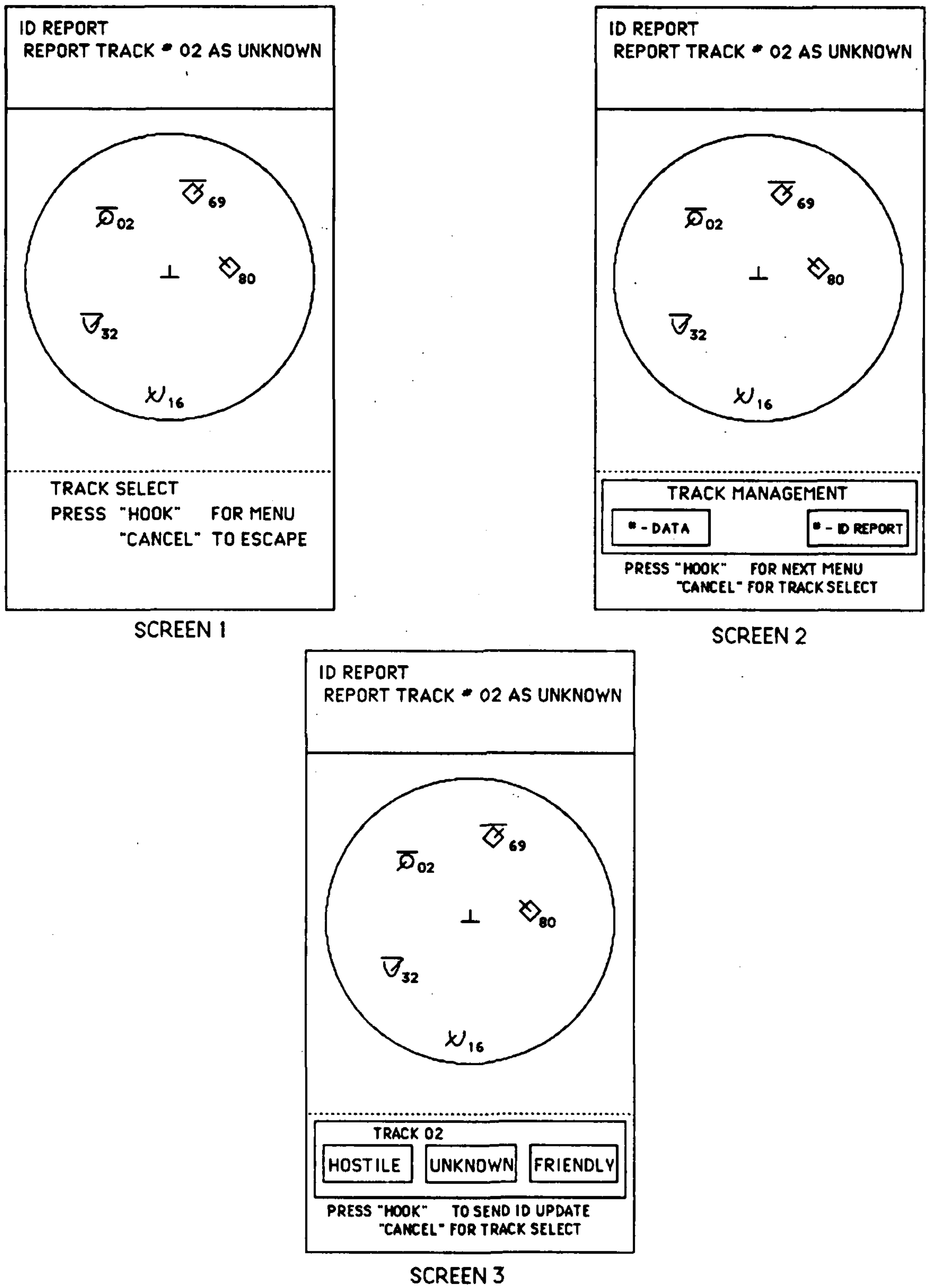


Figure 8. Display screens for identification report task using switches as input modality.

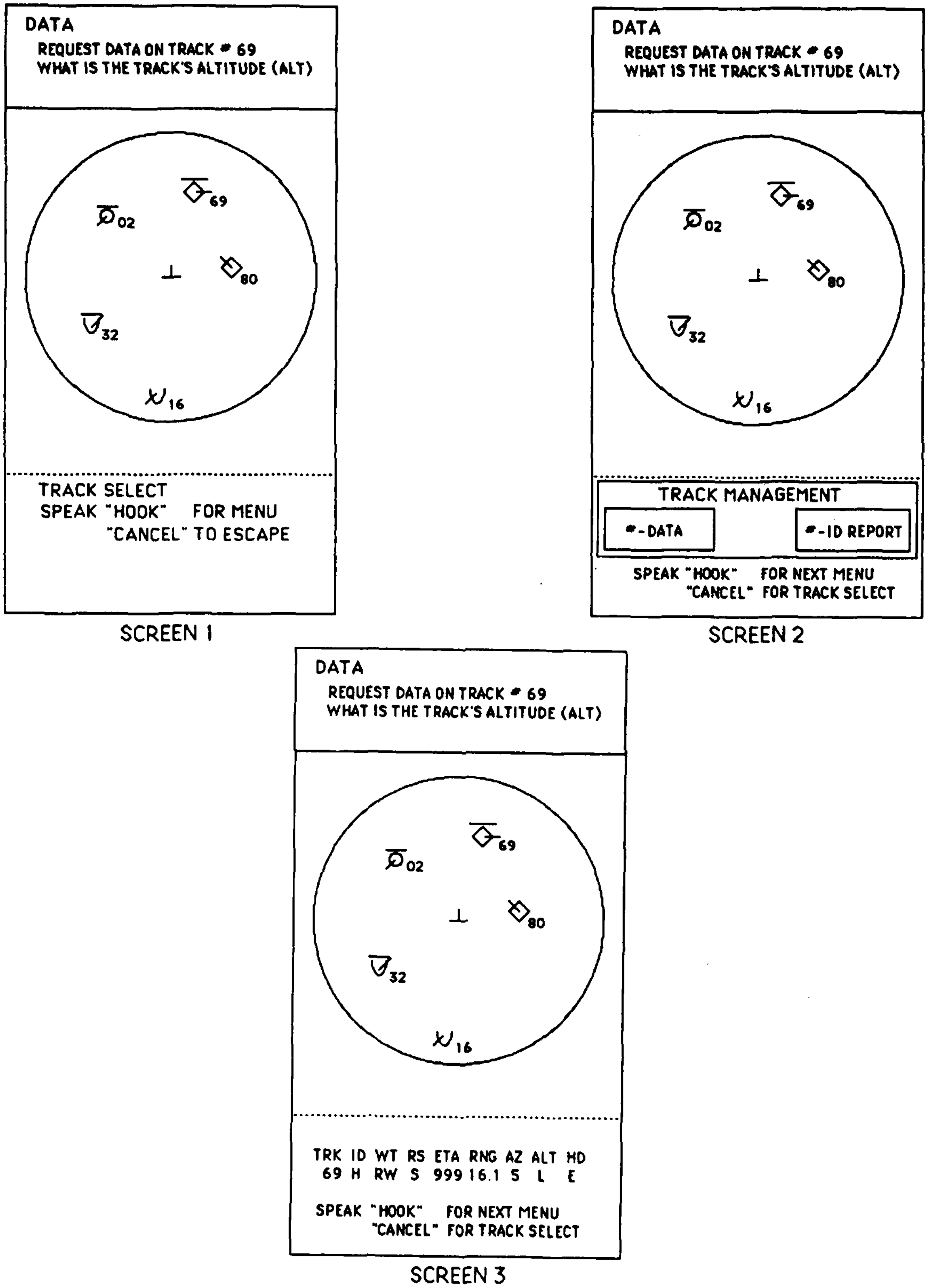


Figure 9. Display screens for data request task using voice as input modality.

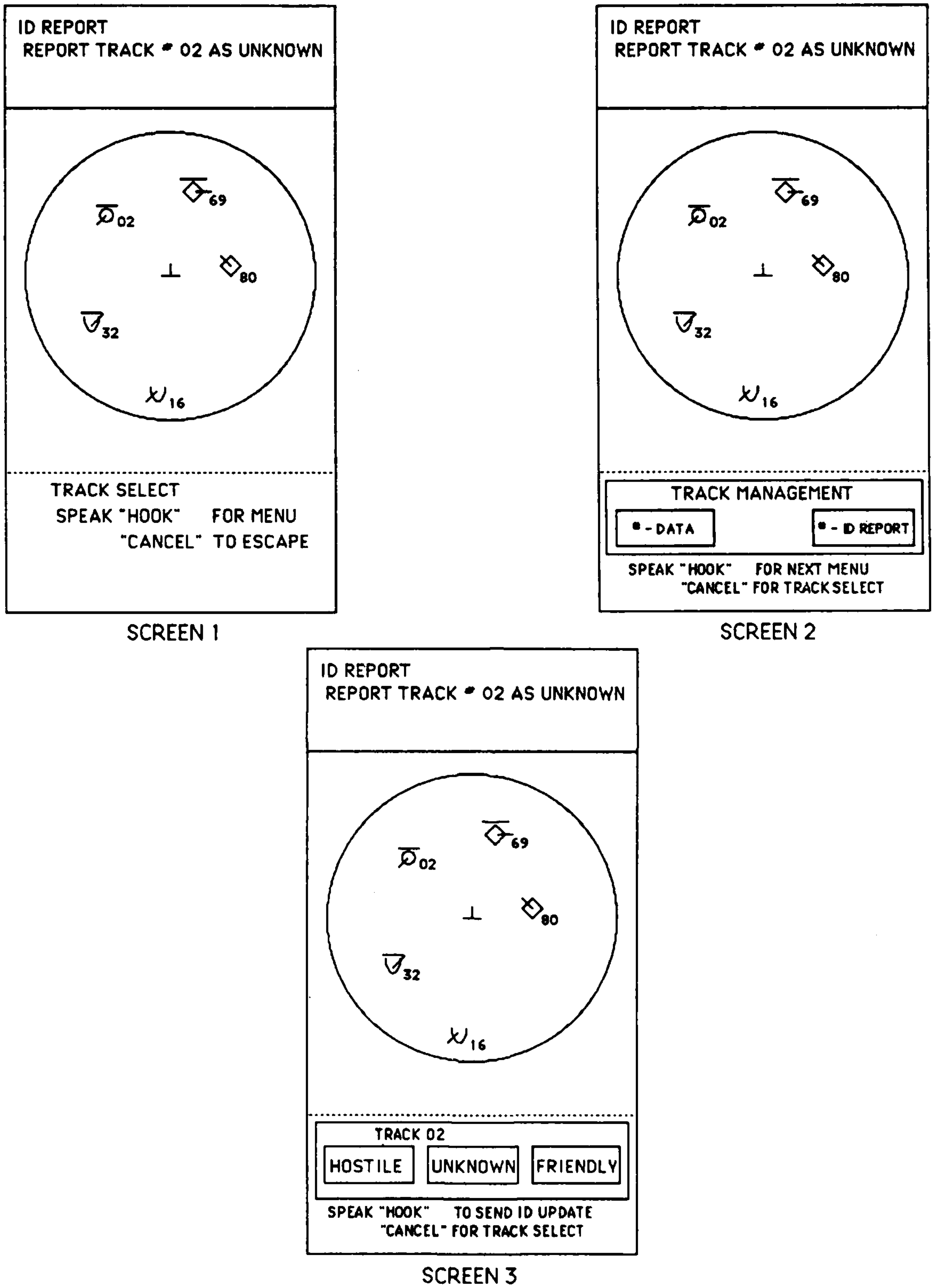


Figure 10. Display screens for identification report task using voice as input modality.

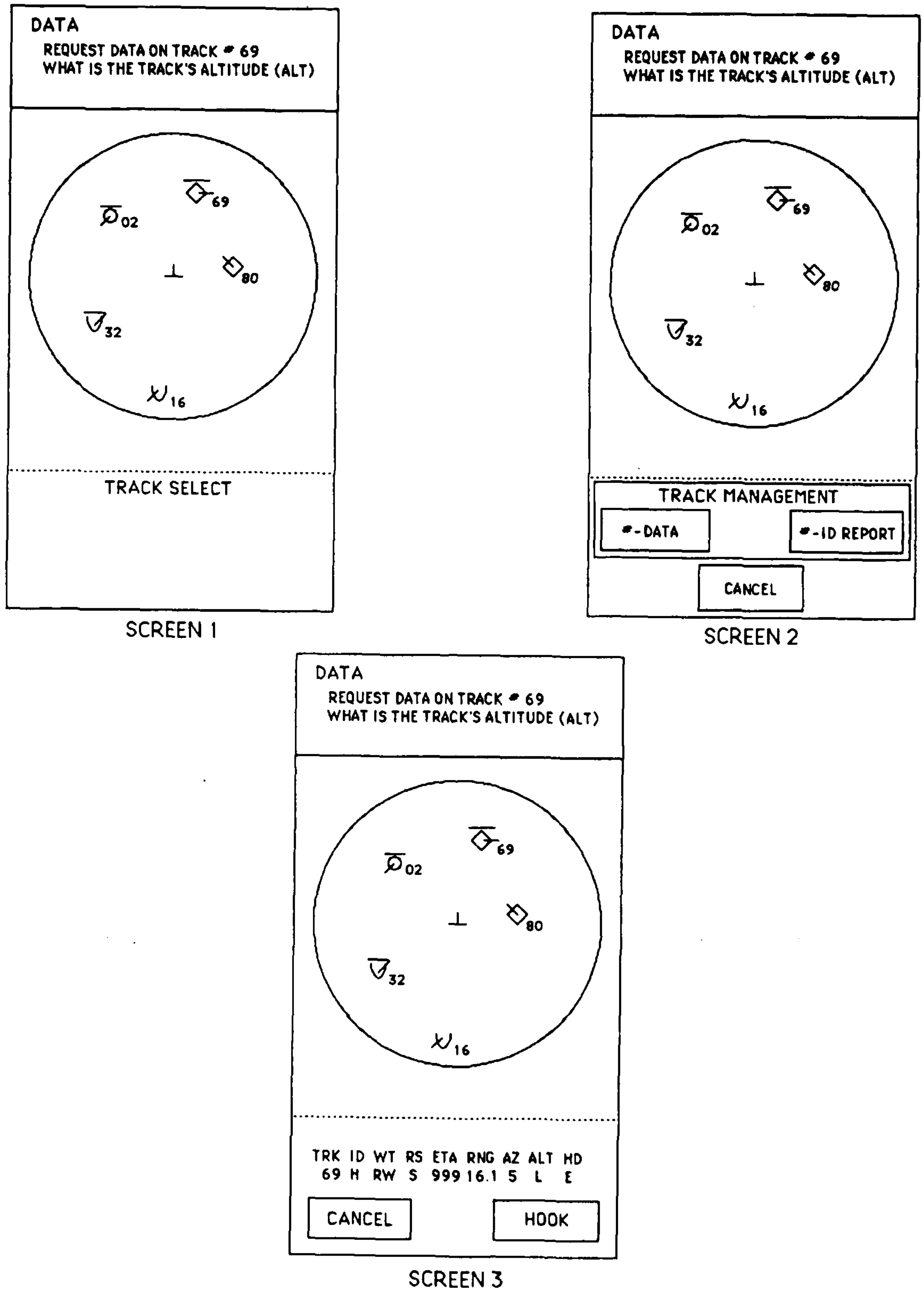


Figure 11. Display screens for data request task using touch panel as input modality.

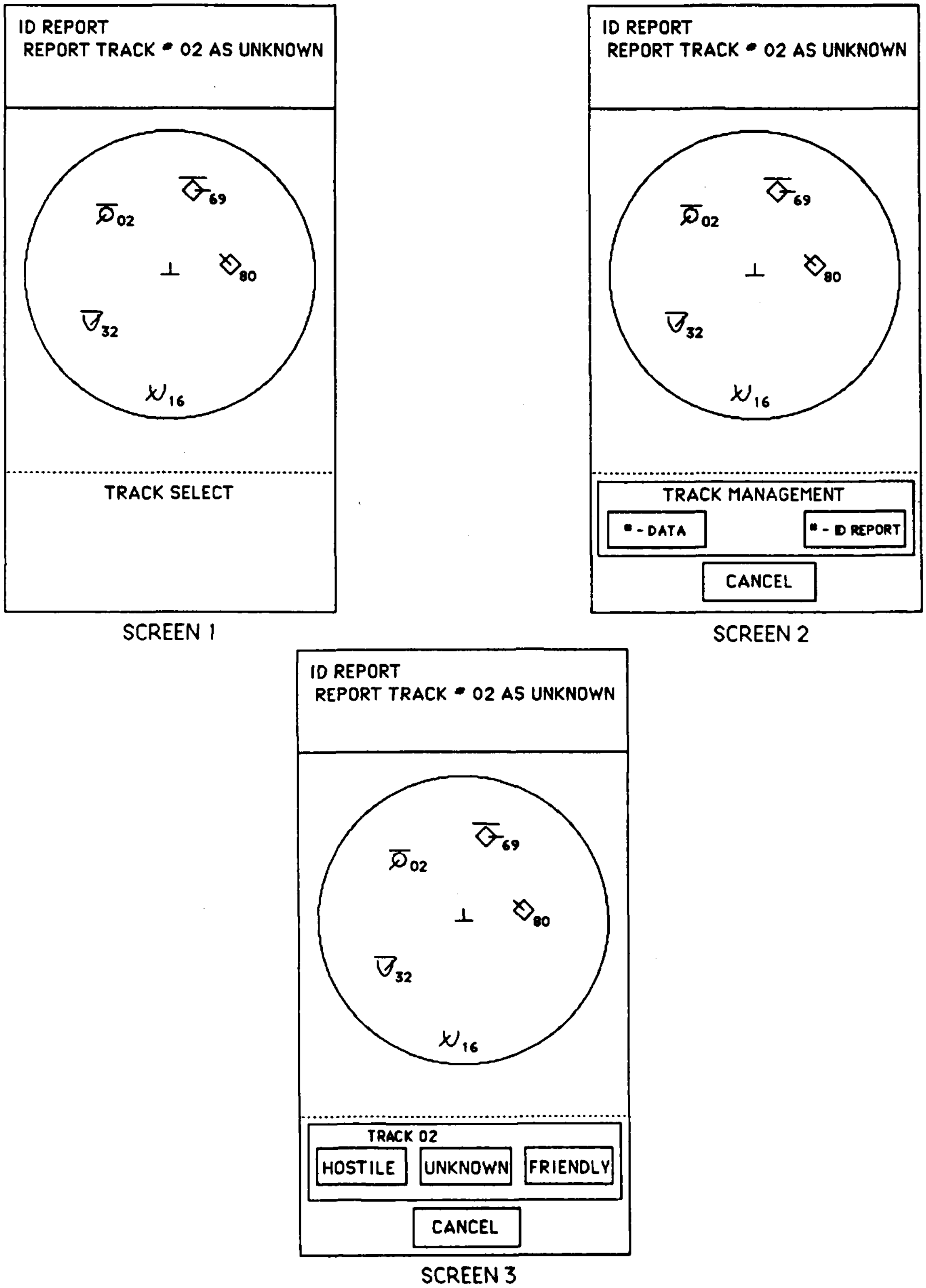


Figure 12. Display screens for identification report task using touch panel as input modality.

show the display presentation for the data request and identification report tasks using switches, voice, and a touch panel as input modalities.

Figures 7 and 8 show the data request and identification report screen formats when using switches as an input modality. Figures 9 and 10 show the data request and identification report screen formats when using voice as an input modality. Figures 11 and 12 show the data request and identification report screen formats when using the touch panel as an input modality.

The instruction area is at the top of the display. Figure 7 shows an example of the data request instruction message, while Figure 8 shows an example of the identification report instruction message used in this investigation. The graphics display area is at the center of the screen, below the instruction area. The tactical graphics area shows a real time scenario with positions of the air tracks updated every second. This area is dedicated to the status of the air picture about the host aircraft and contains five dynamic track symbols. Only one of the track symbols is task-related and has predetermined flight characteristics. A different target track was selected for each task on every trial in a counterbalanced manner so that all flight characteristics were represented on all tasks and input modalities.

The remaining four tracks were selected to simulate live aircraft traffic from a randomly chosen set of parameters in accordance with probabilities listed in Table 2, derived from a study of air defense tactical scenarios (Fallesen, Smyth, and Blackmer, 1983). All flight directions were randomly assigned. The initial positions of the tracks in the scenario were randomly selected with the restriction that no positions be closer than 120 rasters to ensure initial separation allowing display capture (see Display Capture Window section).

The identity of the track was indicated by the symbol shape in accordance with MIL-STD-1477 (Department of Defense, 1983): circular shape for friendly aircraft, diamond shape for hostile aircraft, and U shape for unknown aircraft. Multiple tracks were shown as two symbols, one inside the other. A line above the symbol indicated a rotary wing track; otherwise, the track was a fixed wing aircraft. The track velocity and direction were shown by a track velocity vector. The track designation number (01 to 99) was displayed to the lower right of the symbol.

The menu display area is at the bottom of the screen, below the graphics display area. This area served as the work area for the subject's interaction with the display. A hierarchical menu method of display interaction was chosen (Miller, 1981). In this method, the subject uses a main menu to select submenus from which to work. When the subject is finished with a specific task, he returns to the main menu to repeat the process for the next task. The advantage of using a hierarchical menu approach is that it provides a logical progression of activity while maintaining the task structure. The potential for user disorganization is avoided by using a low number of menu levels (Billingsley, 1982).

The menu area was dedicated to showing the status of the subject's interaction with the display, and listing the menu choices available. The subject's task was limited to information queries about tracks or

Table 2
Scenario Track Probabilities

Factor	Probabilities
Identification	5% Hostile 35% Friendly 60% Unknown
Wing type	35% Fixed wing 65% Rotary wing
Fixed wing	
Velocities	10% At 250 Knots (Kts) 80% At 450 Kts 10% At 450-600 Kts
Altitude	15% Low (0-500 M) 60% Medium (500-4000 M) 25% High (Above 4000 M)
Rotary wing	
Velocity	100% Slow (70-250 Kts)
Altitude	100% Low (0-500 M)
Raid sizes	75% Single 25% More than one

specification of track identification. In general, the copilot-gunner's interaction with the display would be more extensive, including communications about air battle management, battlefield geometry, and command and control messages. For the purposes of this investigation, the assumption is made that the copilot selected a "Track Data Management" option from a master menu. The next logical step would be the choice of the specific track for action, followed by a main menu for the action choice, and a submenu for the action.

The sequence of menu activations on the display for each task was as follows. Figures 10 and 11 illustrate the changes in the menu area.

1. The menu area displays a "TRACK SELECT" prompt. The subject selects the track of interest by one of the input methods to be tested. The completion of a track selection action causes the captured track symbol to blink at a 3-Hz rate and the prompt message to be replaced by the main menu.

2. The main menu lists the two track action options as submenu choices. The two options open to the subject are "DATA" for track data amplification and "ID REPORT" for a change in the track's identification status. The completion of a track action selection causes the main menu to be replaced with the corresponding submenu.

3. The submenu for the track amplification data lists the track number, identification (hostile, unknown, or friendly), wing type (fixed or rotary wing), flight size (single or multiple), estimated time to arrive at host aircraft, range from the host aircraft, azimuth from the host aircraft, altitude (low, medium, or high), and heading (one of the eight cardinal directions). Exiting from the submenu returns the display to the beginning for a new task.

4. The submenu for the track identification task lists the three identification possibilities: friendly, unknown, or hostile. The current identification of the track is highlighted. Selection of the new identification returns the display to the beginning for a new task.

5. In all cases, the option exists for the subject to escape from the main menu or the submenu to the "TRACK SELECT" point by using the "CANCEL" key or voice command. The escape option can be used when the subject selects the wrong track for action, the wrong submenu, or becomes confused and wishes to start the process again. Additionally, all menus and submenus have cue lines to guide the subject in the options available.

h. Display Capture Window. The data selection (capture) window on the display is determined by the spatial tolerance required to operate the different methods. This can be explained as follows. As the subject performs each task, he interacts with the computer by selecting a series of appropriate display items. A display item is selected for processing by the computer when the subject performs a hooking or capturing action by interrupting the infrared beams on the touch panel, or by a switch or verbal entry for the ocular methods. The point of beam interruption for the touch panel or the computed gaze point on the display for the ocular based methods are interpreted as the display point of interest for the subject. The display item closest to the position of the point of interest is selected for processing. The display item must be a member of the subset appropriate for

the subtask: selection of a track symbol, selection of a submenu, or processing a submenu. Furthermore, the display item must be within the display capture window, centered on the point of interest, to preclude responses to accidental selection actions.

The minimum size of the display capture window is determined by the accuracy with which the interest point may be computed for the method. The minimum capture window should be large enough to contain the required spatial tolerance needed for the display activation. The display items, including the tracks in the display scenario, must be far enough apart on the display so that the separation distance between adjacent elements exceeds the spatial tolerance distance. An inaccurate method would require an activation window with a large tolerance, and consequently, a low density scenario with tracks spread far apart. In this test, the window size and scenario densities required to operate the least accurate method were used for all devices. The required tolerance and resulting window size were determined by consideration of the manufacturer's listed accuracies for the different devices comprising each method and verified in a pilot study.

The size of the display window was calculated from the static accuracy of the Polhemus device ($\pm 1.5^{o}$) and the NAC Eyemark oculometer. The accuracy of the NAC oculometer is a function of the accuracy of several variables: (a) face mask stability, (b) calibration, (c) sensor resolution, (d) corneal reflection detection, and (e) differences in anatomical shapes of the eyeball and orbit among subjects. The overall accuracy of the Eyemark is reported to be about 1.8^{o} assuming a stable face mask (NAC, 1987). The combined accuracy for the two independent devices is $\pm 2.34^{o}$. This corresponds to a $\pm$1.1-inch display accuracy at the nominal 27-inch viewing distance. The display is 1024 rasters by 1024 rasters on an 11-inch by 13-inch monitor, with 72.72 rasters per inch in the x-direction (vertical) and 95.17 rasters per inch in the y-direction. The computed display accuracy translates to a 160- by 210-raster display capture window centered on the predicted gaze point. The corresponding track minimum separation distance is 210 rasters.

Test Subjects

The test subjects were 15 right-handed, male, military enlisted personnel of rank E-6 or below. All were required to be right-handed to preclude confounding subject handedness and the keypad position on the console. Demographic data collected for each test subject are presented in Table 3. All 15 subjects were assigned to the Field Support Branch of the U.S. Army Combat Systems Test Activity (USACSTA), Aberdeen Proving Ground, Maryland, and had prior experience as test subjects on military systems.

Experimental Design

The experiment was a repeated measures (15x5x2) design with tasks and methods within subject fixed factors, and subjects as a random variable. The fixed factors are the (a) data entry method (five levels: oculometer with switch, oculometer with voice, fixed reticle with switch, fixed reticle with voice, and touch panel) and (b) task type (two levels: track data report and track identification report). The touch panel was included to allow the comparison of the performance of the other devices to that of a standard data

Table 3
Test Subject Demographic Data

Subject	Rank	Military occupational specialty (MOS)	Date of birth (DOB)	Education (Years)
1	E-4	11H10	020254	14
2	E-5	11B10P	073064	13
3	E-5	11M10	011059	13
4	E-3	11M10	102164	12
5	E-4	11B	041167	12
6	E-2	88M10	121168	12
7	E-4	88M	032166	12
8	E-4	88M10	072661	12
9	E-4	19K10R	080757	13
10	E-4	12F10	040167	12
11	E-5	19KR2	101958	16
12	E-5	19E2R	110161	12
13	E-6	19K3R	122655	12
14	E-4	12B	040158	12
15	E-4	12B10	072365	12

entry device. The dependent variables were the task times and the number of task errors. The task times were the total time to complete the tasks, and the component times: (a) time to select a track symbol, (b) time to select a submenu, and (c) time to complete a submenu action. The task errors are the data input errors made while performing the tasks.

The data collected during the test included the data entries made by the subject and the times at which the entries were made. The time to perform the task and the component times were computed from the recorded times. The errors were determined by comparing the entries made to what should have been entered.

Training and Test Procedures

Each subject was trained and tested individually and all were given the same training. The training and testing for each subject in this experiment were separated into three consecutive sessions lasting no more than 4 hours total. The three sessions were (a) instruction, (b) orientation, and (c) training and testing the subjects. The first session consisted of detailed explanations of the tasks and the data entry methods. Each test subject read an explanation of the study and was given the opportunity to ask questions. Diagrams of the displays at each level of menu interaction were used to help explain the tasks.

In the second session, the subject was operationally familiarized with each of the data entry methods in the following fixed order:

1. Touch panel
2. Fixed reticle with switch
3. Fixed reticle with voice
4. Oculometer with switch
5. Oculometer with voice

Every subject was allowed to operate each method for two training runs. The purpose of this phase was to familiarize the subject with the methods and test procedures before training and testing, since some of the methods were new and may have never been seen before by the subject.

The subject was then trained and tested about each of the five data entry methods in turn according to an assigned test sequence. The training and testing began with the calibration of the oculometer or reticle or enrollment of the speech recognizer, as appropriate.

In the calibration process for the fixed reticle method, the subject aligns the reticle cross hairs with cue marks on the display face. The appropriate relations among the sensor coordinates and the position of the subject's eye and viewing direction through the reticle, are computed using a least squares regression analysis technique from the known locations of the

cue marks in the 2-dimensional coordinate system of the display, the 3Space sensor returns for each of the sightings and the measured distances from the sensor to the reticle cross hair.

The fixed reticle used in this test was a cross hair engraved on the plastic visor. When in place, the reticle was about 2 inches in front of the subject's preferred eye. This is close enough so that the device is not in the vision field of the other eye. The reticle is not projected to infinity, and appears as a blurred dark image that is visible against the lighted background of the display. Following adjustment of the screen brightness to a level between that necessary to accommodate display legibility and the visibility of the cross hair, the subject quickly learned to center the cross hair pattern over the image of the display items.

The Eyemark recorder must be centered and aligned to the subject's vision field for accurate results with the oculometer method. The device is centered on the subject's head, and the mirrors and focus are adjusted to ensure that the corneal reflection is within the field of view of the tracking camera units as shown by the eye-pupil image on the television monitor. The subject is then directed to gaze at the center of a calibration pattern, keeping his head in a fixed position. The experimenter adjusts the camera mirrors so that the Eyemark returns agree with the pattern shown on the television monitor. Finally, the subject is directed to gaze at each of the extremes of the pattern in turn, keeping his head fixed, as the experimenter adjusts the magnification of the camera return to agree with the calibration pattern.

The first calibration step for the oculometer method uses the fixed reticle to establish the relation between the subject's eye position and the sensor as in the fixed reticle method. The second calibration process establishes the relation between the oculometer coordinate axes and those of the sensor. In the second process, while in a fixed head position, the subject looks at a series of cue marks on the display. The appropriate relations are computed using a least squares regression analysis technique from the known locations of the cue marks in the 2-dimensional coordinate system of the display, the Eyemark returns and the 3Space sensor returns for each of the sightings, as well as the computed eye position relative to the sensor.

The accuracy of the oculometer depends on proper alignment and stability of the head support system which is needed to maintain boresight. A 0.1-millimeter shift in mask position on the face will cause a 1.4° shift in boresight (NAC, 1987). The mask is easily shifted on the face by changes in head position and expression. The 1.6 pounds (720 grams) weight of the mask is to the front, and quick vertical movements in head position cause slight shifts. These mask shifts are reduced by maintaining a fixed head position, however. For these reasons, the subject was instructed to keep his head fixed in position and oriented at a predetermined reference point during alignment and calibration, and during testing to move his head as little as possible. The inclusion of a calibration check between test trials allowed the experimenters to check the calibration and correct for drifts in alignment.

The field of view (FOV) of the oculometer is 45° by 60°. The display subtends less than 21° by 27° (10 inches by 13 inches at 27 inches' viewing distance). The entire display is easily seen without head movements by shifting one's eyes. Subjects quickly learned to operate the oculometer from

this position. There is evidence that people naturally shift their eyes without head movements to view items out to 12^{o} off-center. This span covers all pertinent data items on the display.

Athough the static accuracy of the Polhemus is 1.5^{o} over a wide range of head movements, the subject was asked to maintain his head at the calibration point to ensure a constant viewing distance and consistent test procedure. Subjects were able to maintain the reference position without difficulty throughout the test.

Figure 13 shows the subject during the oculometer alignment and associated calibration processes. Figure 14 is a view of the calibration pattern which the subject sees on the Aydin display.

The subject must be enrolled on the automatic speech recognizer before the device can be used in applications. The enrollment process establishes reference templates for each of the command words. The command words are downloaded to the VRM for the start of the enrollment period. The prompts from the VRM are shown on the display monitor as a guide for the subject. A built-in 1-second delay between prompts following subject verbal response ensures isolated word response. The enrollment is conducted in a noise-free environment.

Following the calibration and enrollment, the subject was instructed in hands-on training for the data entry method to be tested. Each test subject completed a total of ten training and five test trials for each data entry method. A trial was comprised of two successive tasks: a track data request task and a track identification update task, assigned in random order. Each task started with an instruction to the subject to perform the appropriate data entry or data extraction actions for a specific track. The subject read the instruction prompt and selected the track by hooking the track symbol from the display with the data entry technique being tested and tagging it for further processing. Next, the subject selected the submenu appropriate for the instructed task. The submenu appeared on the display and the subject completed the task. A calibration pattern (Figure 14) was displayed between trial runs. The pattern was used for the oculometer methods to allow the experimenter to check the alignment before starting the next test run. Figure 15 shows the subject during such a calibration check for the oculometer conducted between test trials to ensure that the subject has remained within calibration. Upon completion of the training and testing of a particular configuration, the subject proceeded to the next data entry method in the assigned test sequence, until all five methods were tested. The assignment of testing sequences was counterbalanced among test subjects. The counterbalancing scheme is listed in Table 4.

Each test subject was given a posttest debriefing after completing the test. The questionnaire administered to each test subject is presented in the Appendix.

RESULTS

The results of the statistical analysis of the objective data and of the subjective survey are covered in this section.

Figure 13. Calibration of oculometer.

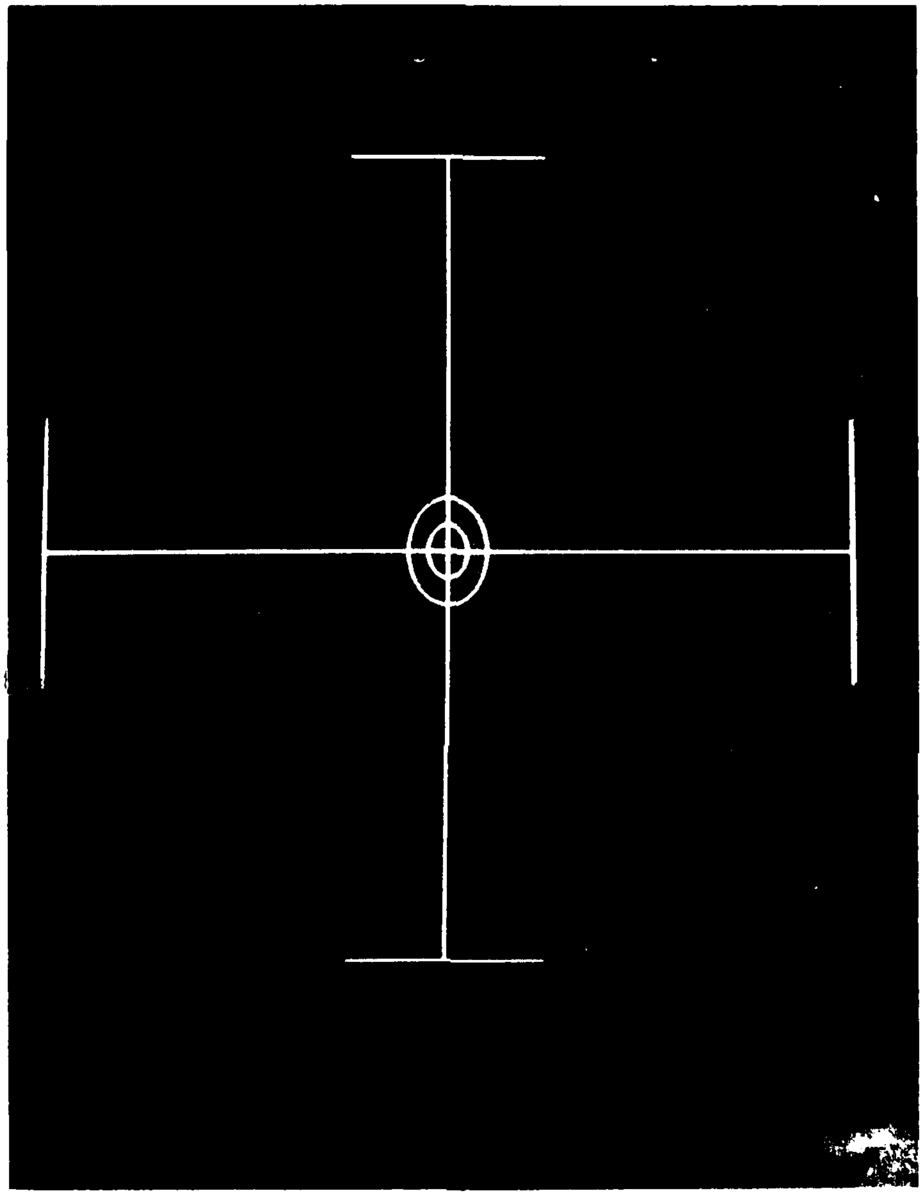


Figure 14. Calibration pattern for oculometer.

Figure 15. Calibration check of oculometer.

Table 4
Testing Sequence Assignments

Subjects	First	Second	Third	Fourth	Fifth
1, 6, 11	OS	OV	FS	FV	TP
2, 7, 12	OV	TP	FV	OS	FS
3, 8, 13	FS	OS	TP	OV	FV
4, 9, 14	FV	FS	OS	TP	OV
5, 10, 15	TP	FV	OV	FS	OS

KEY

TP - TOUCH PANEL
FV - FIXED RETICLE WITH VOICE ENTRY
FS - FIXED RETICLE WITH SWITCH ENTRY
OV - OCULOMETER WITH VOICE ENTRY
OS - OCULOMETER WITH SWITCH ENTRY

Objective Data

A multivariant analysis of variance (MANOVA) was used for the statistical analysis of the transformed values of the four time variables: (a) time to hook target symbol, (b) time to select submenu, (c) time to complete action on submenu, and (d) the total time to perform a task, which is the sum of the above three times. The time data were transformed to their inverse to generate a normal distribution and to reduce the correlation between mean and variance, conditions necessary for the valid application of parametric statistical analysis. The data for each variable were summed across replications to obtain one sample per subject for each task by method combination. The statistical analysis of the errors was a chi-squared nonparametric test appropriate for count data.

Table 5 lists the results for the repeated measures MANOVA of the transformed dependent time variables as measured by Wilks' criterion. Table 5 also lists Mauchly's criteria for sphericity as a test of compound symmetry for repeated measures data, the correlation between the means and variances of the task by method combinations for the transformed time variables, and a test value that the corresponding distributions are not from a normally distributed population. All analysis was done using the SAS Statistical Analysis Computer Package on a DEC VAX 11/780.

The listings in Table 5 show that the multivariant analysis is statistically significant at the 0.05 confidence level, as measured by the Wilks' criterion, for the methods and tasks for all time variables. The method by task interaction is statistically significant for the time to complete the subtask (T3), but not for the other time variables. The assumption of compound symmetry is mainly satisfied for the time variables. The means of each task by method cell are low to moderately correlated with the corresponding variances. Finally, the median statistic listed for the normality assumption applied to the distributions of the task by method combinations suggest that this assumption is not violated.

The results of the corresponding univariate analysis of variance for within-subject effects are summarized in Table 6 for each of the dependent time variables. The Greenhouse-Geisser corrections for departure from compound symmetry are slight in keeping with the insignificant values for Mauchly's criteria. The results of the analysis agree with those from the multivariant analysis.

The results of the Scheffē pair-wise comparison test applied to the dependent variables for the methods are listed in Table 7. The average total task time for the reticle switch is significantly faster than the oculometer voice. The times for the touch panel, the oculometer with switch, and the reticle and voice are not significantly different from either the time for fixed reticle and switch or oculometer with voice.

The chi-square nonparametric test applied to the number of total errors shows a significant difference between methods at the .05 confidence level. Table 8 lists the observation matrix for the number of total errors as a function of the joint occurrences of the method and task. Table 8 lists the chi-squared computations for the task, method, and task by method

Table 5

Results of the Repeated Measures Multivariant Analysis of Variance for the Transformed Time Variables: Wilks' Criteria and Mauchly's Criteria for Sphericity, the Correlation of the Means and Variances, and the Normality Test

Effect	Wilks' criteria	Mauchly's criteria	Correlation	Normality test
T1: Time to select track				
Task (T)	0.0004	-		
Method (M)	0.0107	0.5142		
T X M	0.2941	0.5132	0.0164	0.346
T2: Time to select submenu				
Task (T)	0.0368	-		
Method (M)	0.0018	0.5578		
T X M	0.0702	0.7100	0.0759	0.616
T3: Time to complete submenu				
Task (T)	0.0001	-		
Method (M)	0.0001	0.2758		
T X M	0.0315	0.2862	0.2292	0.526
TT: Total time to complete task				
Task (T)	0.0001	-		
Method (M)	0.0002	0.5221		
T X M	0.5707	0.4142	0.6533	0.520

KEY

- T1 - TIME TO SELECT TRACK SYMBOL
- T2 - TIME TO SELECT SUBMENU
- T3 - TIME TO COMPLETE SUBMENU
- TT - TOTAL TASK TIME

Table 6

Univariate Analysis of Variances for Within-Subject Effects of the Transformed Time Variables

Effect	df	MS	F	Pr > F	Adjusted[a] Pr > F
I. T1: Time to select track					
Task	1	0.0213	21.17	0.0004	-
Error	14	0.0010			
Method	4	0.0157	5.11	0.0014	0.0043
Error	56	0.0031			
Task X Method	4	0.0011	1.40	0.2468	
Error	56	0.0008			
II. T2: Time to select submenu					
Task	1	0.0012	5.32	0.0368	-
Error	14	0.0002			
Method	4	0.0268	7.76	0.0001	0.0002
Error	56	0.0034			
Task X Method	4	0.0016	2.89	0.0302	0.0388
Error	56	0.0005			
III. T3: Time to complete subtask					
Task	1	0.4984	635.08	0.0001	
Error	14	0.0007			
Method	4	0.0382	22.35	0.0001	
Error	56	0.0017			
Task X Method	4	0.0054	6.75	0.0009	
Error	56	0.0008			
IV. TT: Total time to complete task					
Task	1	0.0563	261.95	0.0001	-
Error	14	0.0002			
Method	4	0.0172	12.33	0.0001	0.0001
Error	56	0.0014			
Task X Method	4	0.0003	1.41	0.2430	0.2553
Error	56	0.0002			

[a] Adjusted by Geisser-Greenhouse correction for sphericity.

KEY
T1 - TIME TO SELECT TRACK SYMBOL
T2 - TIME TO SELECT SUBMENU
T3 - TIME TO COMPLETE SUBMENU
TT - TOTAL TASK TIME

Table 7
Scheffé's Test for the Methods

	Scheffé grouping		Mean time (seconds)	Number	Method
Total time to complete task		A	17.4930	150	OV
	B	A	14.9530	150	FV
	B	C	14.0300	150	OS
	B	C	12.7560	150	TP
		C	12.2670	150	FS
Time to select target		A	8.0064	150	OV
	B	A	6.4537	150	FV
	B	A	6.2927	150	OS
	B		5.6221	150	FS
	B		5.4784	150	TP
Time to select submenu		A	4.3235	150	OV
	B	A	3.8298	150	OS
	B	A	3.7475	150	FV
	B		3.2727	150	TP
	B		3.1285	150	FS
Time to complete submenu		A	5.1628	150	OV
		A	4.7517	150	FV
	B		4.0046	150	TP
	B		3.9074	150	OS
	B		3.5162	150	FS

Note: Means with same grouping letter are not significantly different at the .05 level.

KEY

- TP - TOUCH PANEL
- FV - FIXED RETICLE WITH VOICE ENTRY
- FS - FIXED RETICLE WITH SWITCH ENTRY
- OV - OCULOMETER WITH VOICE ENTRY
- OS - OCULOMETER WITH SWITCH ENTRY

Table 8

Chi-Squared Analysis of the Total Errors by Task, Method, and Task by Method Interaction (0.05 Significance Level)

Observation Matrix

Method	**Task ID**	**Task DA**	**Total**
TP	2	1	3
FV	25	18	43
FS	21	25	46
OV	43	55	98
OS	36	35	71
Total	127	134	261

(a) Task by method interaction

df	Chi-squared	Test statistic
4	3.12	9.49

(b) Method

df	Chi-squared	Test statistic
4	95.68 [a]	9.49

(c) Task

df	Chi-squared	Test statistic
1	0.18	3.84

[a] Significance at the .05 confidence level.

KEY	
METHOD	TP - TOUCH PANEL
	FV - FIXED RETICLE WITH VOICE ENTRY
	FS - FIXED RETICLE WITH SWITCH ENTRY
	OV - OCULOMETER WITH VOICE ENTRY
	OS - OCULOMETER WITH SWITCH ENTRY
TASK	DA - DATA REQUEST
	ID - IDENTIFICATION CHANGE

interactions. The computations show that the effects of the methods and tasks on the errors are independent. Similarly, the effect of tasks on errors is insignificant.

Table 9 lists the number of samples, averages, standard deviations and standard error statistics for the different task times and the total errors as a function of the method by task interaction.

Subjective Data

Subjects ranked the five methods by their preference on a 5-point scale from 1 (most preferred) to 5 (least preferred). Table 10 lists the rankings and shows that the results are mixed. The rankings were statistically analyzed by the Friedman two-way analysis of variance by ranks for related samples, a nonparametric test. The rankings are independent of method at the 0.05 level of significance, since the computed Friedman statistic is less than the chi-square test statistic. Table 10 is a summary. The results of the cross-tabulations of the ratings showing the subject preferences are listed in Table 1 in the Appendix. The posttest debriefing questionnaire and the frequency of additional comments by the test subjects are listed also in the Appendix.

DISCUSSION

In this portion of the report, we discuss the experimental results, and particularly, the average task times, the task and method interaction for the complete menu subtask times, the errors and the relation of the times to the errors, sources of selection errors, speech recognizer errors, the dropped subjects, and the nonerror times and spatial tolerance.

Average Task Times

The task times are determined by the different activities unique to operating the data entry methods. For all methods, the response of the subject at the start of the trial to the instruction line is to first visually scan the display for the track of interest and then select the located track. The selection is made for the touch panel and switch-based methods by the mechanical action of a person moving his hand to the display item for the touch panel or pushing the switch while gazing at the item of interest for the switch methods. An intermediate step is required for the fixed reticle during which the subject moves his head slightly to align the reticle with the image of the display item. The head movements required are slight and take less than a second (List, 1983). The voice-based methods take longer because of the additional time to utter the appropriate verbal command and the time for the speech recognizer to process the utterance before data transfer to the host computer.

The rank ordering of the task times agrees with this discussion. For example, the rank ordering of the average total time to complete a task, from quickest to slowest, of the five methods (see Table 7) is the (a) reticle switch, (b) touch panel, (c) oculometer switch, (d) reticle voice, and (e)

Table 9
Descriptive Statistics for the Task Times and Total Errors for the Method by Task Interaction: Number, Average, Standard Deviation, and Standard Error

Factor (device & task)		Number	Average	Std. dev	Std. error
Total task time (seconds)					
Method	Task				
TP	DA	75	13.7997	3.8670	0.4465
	ID	75	11.7117	3.1387	0.3624
FS	DA	75	13.1565	3.7767	0.4361
	ID	75	11.3772	3.9826	0.4599
FV	DA	75	16.5337	4.2276	0.4882
	ID	75	13.3720	5.3516	0.6179
OS	DA	75	14.9793	4.8790	0.5634
	ID	75	13.0805	7.0034	0.8087
OV	DA	75	19.6827	10.3282	1.1926
	ID	75	15.3028	8.5643	0.9889
Time to select target (seconds)					
Method	Task				
TP	DA	75	5.6783	2.1875	0.2526
	ID	75	5.2784	1.8229	0.2105
FS	DA	75	5.7258	2.3868	0.2756
	ID	75	5.5184	2.7640	0.3192
FV	DA	75	6.5198	2.4150	0.2789
	ID	75	6.3875	4.1409	0.4781
OS	DA	75	6.6166	3.3582	0.3878
	ID	75	5.9688	3.4996	0.4041
OV	DA	75	8.9648	8.8867	1.0261
	ID	75	7.0479	7.2016	0.8316
Time to select submenu (seconds)					
Method	TASK				
TP	DA	75	3.5004	1.4998	0.1732
	ID	75	3.0450	0.9283	0.1072
FS	DA	75	3.1455	1.0620	0.1226
	ID	75	3.1115	1.1117	0.1284
FV	DA	75	3.8441	0.9436	0.1090
	ID	75	3.6509	1.0190	0.1177
OS	DA	75	3.6601	1.8877	0.2180
	ID	75	3.9996	3.3748	0.3897
OV	DA	75	4.2183	1.6941	0.1956
	ID	75	4.4288	1.7277	0.1995

(Continued)

Table 9 (Continued)

Factor (device & task)		Number	Average	Std. dev	Std. error
Time to complete subtask (seconds)					
Method	**Task**				
TP	DA	75	4.6209	1.3169	0.1521
	ID	75	3.3884	1.6958	0.1958
FS	DA	75	4.2852	1.4042	0.1621
	ID	75	2.7473	0.9455	0.1092
FV	DA	75	6.1697	1.9702	0.2275
	ID	75	3.3336	0.9295	0.1073
OS	DA	75	4.7025	1.4376	0.1660
	ID	75	3.1122	1.1047	0.1276
OV	DA	75	6.4995	1.7047	0.1968
	ID	75	3.8261	1.9303	0.2229
Total errors					
Method	**Task**				
TP	DA	75	0.0133	0.0000	0.0000
	ID	75	0.0266	0.1622	0.0187
FS	DA	75	0.3333	0.6600	0.0762
	ID	75	0.2800	0.7035	0.0812
FV	DA	75	0.2400	0.5123	0.0591
	ID	75	0.3333	0.7542	0.0871
OS	DA	75	0.4667	0.9141	0.1055
	ID	75	0.4800	0.9981	0.1153
OV	DA	75	0.7333	1.4079	0.1626
	ID	75	0.5733	1.1682	0.1349

KEY	
METHOD	TP - TOUCH PANEL
	FV - FIXED RETICLE WITH VOICE ENTRY
	FS - FIXED RETICLE WITH SWITCH ENTRY
	OV - OCULOMETER WITH VOICE ENTRY
	OS - OCULOMETER WITH SWITCH ENTRY
TASK	DA - DATA REQUEST
	ID - IDENTIFICATION CHANGE

Table 10
Statistical Analysis of the Subjects' Preference Ratings by the Friedman Two-Way Analysis of Variance for Related Samples

(a) Preference Rating of Subjects as a Function of Methods (Rating: 1 - Most Preferred, 5 - Least Preferred)

	Method				
Subject	TP	FV	FS	OV	OS
1	1	3	4	5	2
2	2	1	3	5	4
3	1	5	2	4	3
4	5	2	4	1	3
5	1	5	4	3	2
6	3	1	2	4	5
7	5	3	1	2	4
8	5	1	2	3	4
9	1	3	2	5	4
10	5	1	2	3	4
11	3	2	1	5	4
12	1	3	4	2	5
13	5	2	4	1	3
14	1	2	3	4	5
15	3	2	1	5	4
Total	42	36	39	52	56

(b) Summary Table for the Friedman Two-Way Analysis of Variance by Ranks for Related Samples (0.05 Significance Level)

N	*k*	*df*	Friedman statistic	Chi-square test statistic
15	5	4	7.893	9.49

KEY
TP - TOUCH PANEL
FV - FIXED RETICLE WITH VOICE ENTRY
FS - FIXED RETICLE WITH SWITCH ENTRY
OV - OCULOMETER WITH VOICE ENTRY
OS - OCULOMETER WITH SWITCH ENTRY

oculometer voice. The touch panel, oculometer switch, and reticle voice are statistically grouped together. The switch methods are faster than the voice, and the fixed reticle is faster than the oculometer.

Task and Method Interactions

The results of the repeated measures analysis listed in Tables 5 and 6 show that the task by method interaction is significant (0.05 confidence level) for the time to complete the submenu (T3). Figures 16 through 19 are plots of the average submenu times for each of the five methods performing the two tasks. The plot shows that the voice methods required longer than the switch methods and the touch panel when performing the data request task (see Figure 18); the performance was about the same for all the methods on the identity specification task, however. The difference in times on the data request task is about 1.75 seconds. The subjects had to pause after reading the requested data to the experimenter before voicing the "exit" command to the speech recognizer. Since the recognizer processes connected speech but not continuous speech, the pause was necessary to prevent running the two utterances together.

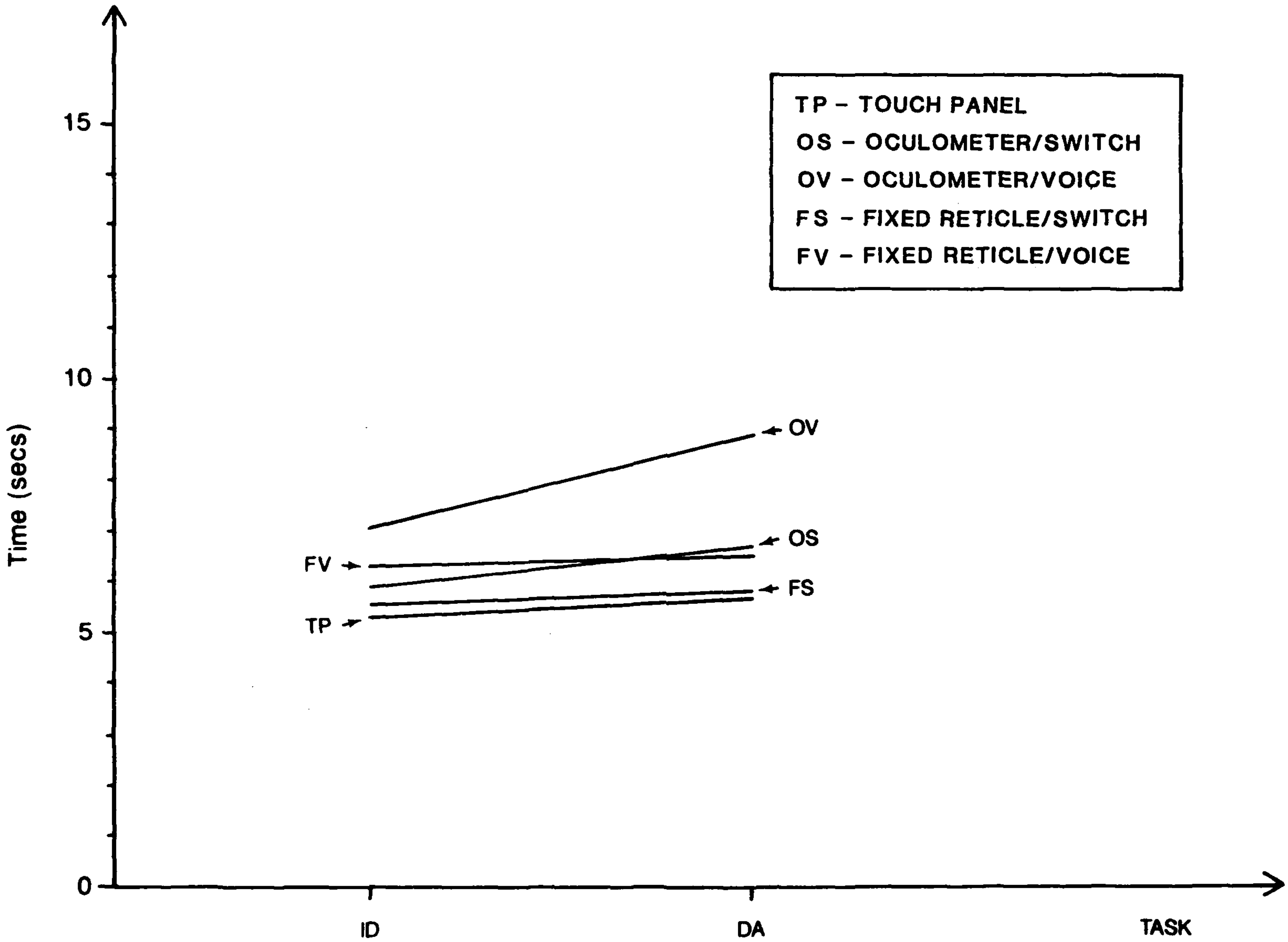
Task Errors

The task errors are caused by the subjects selecting incorrect items inside the display capture window (see Methodology, Apparatus section), or by trying to select display items outside the display window. The subject may make these errors in one of two ways. First, he may make a cognitive mistake in selecting the wrong track or subtask in disagreement with the instruction. He then either completes the task erroneously or selects the "cancel" action to start the task again. Another source of error is mechanical misalignment of the selection point with the display item. The attempt to select a display item outside the display window will result in a hook action error if no other item is within the window. Again, the selection of a wrong display item may be dropped with a cancel action to start the task again, or the subject may erroneously process it to completion.

The results section shows that the total number of errors differ significantly by method, but not by task or method and task interaction. The different types of errors made by the subjects are listed in Table 11 as "completed" errors and "selection" errors as a function of method. Completed errors are not corrected before the completion of the task. A selection error is generated by any hook action exceeding the three selection actions required for correct performance of the task, and by any cancel actions needed to correct erroneous hooking actions. These errors are a measure of the subject's excessive activity in performing his task.

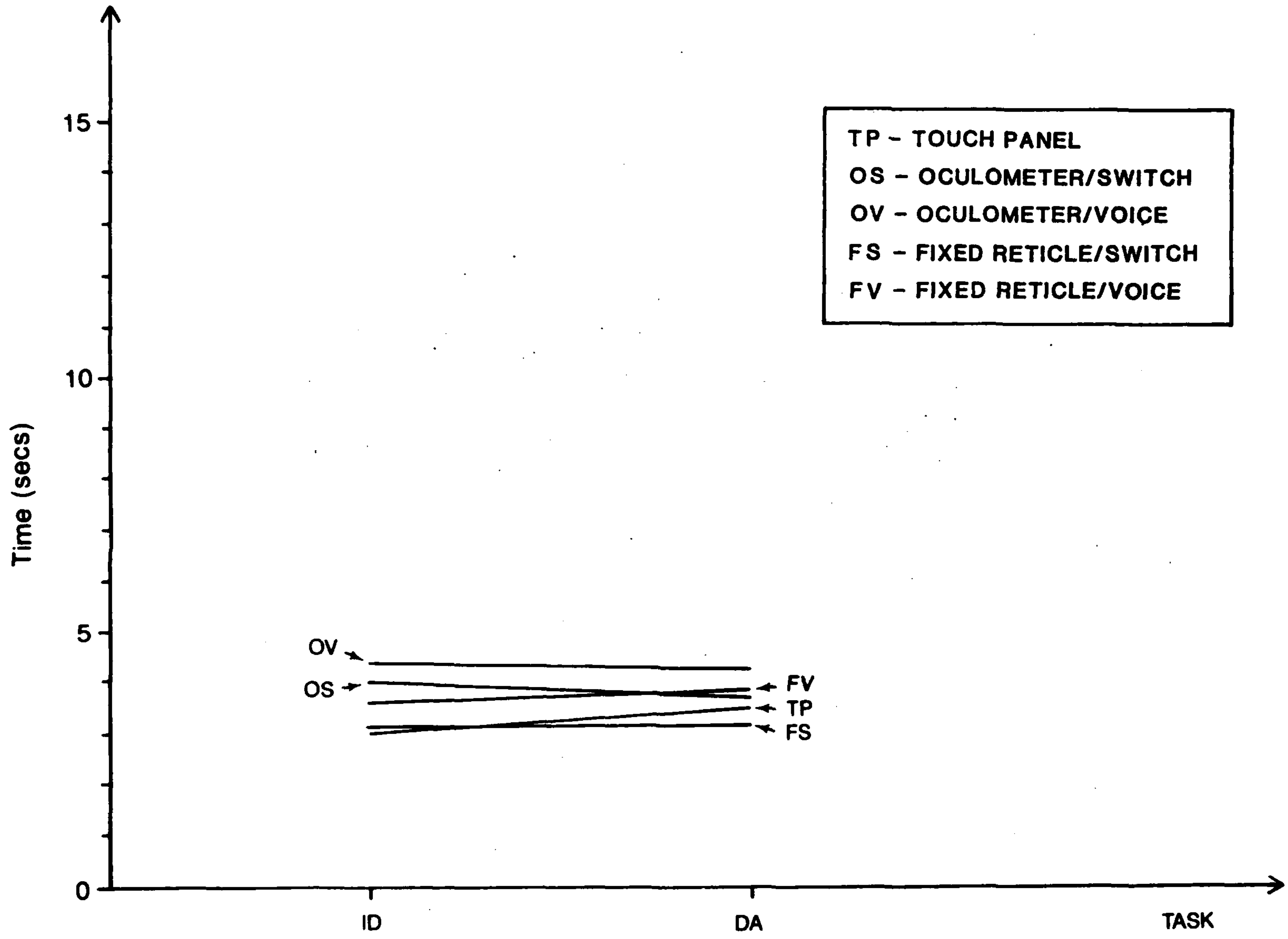
The total number of errors is 264 for the 750 test runs conducted in this experiment. On the average, this is one error every third test run. The total number of errors are the sum of the number of completed errors and the number of selection errors.

Table 11 shows that very few completed errors were made. Most of the 31 completed errors were caused by processing wrong tracks (24 errors). Three errors were caused by selecting the wrong track identification on the



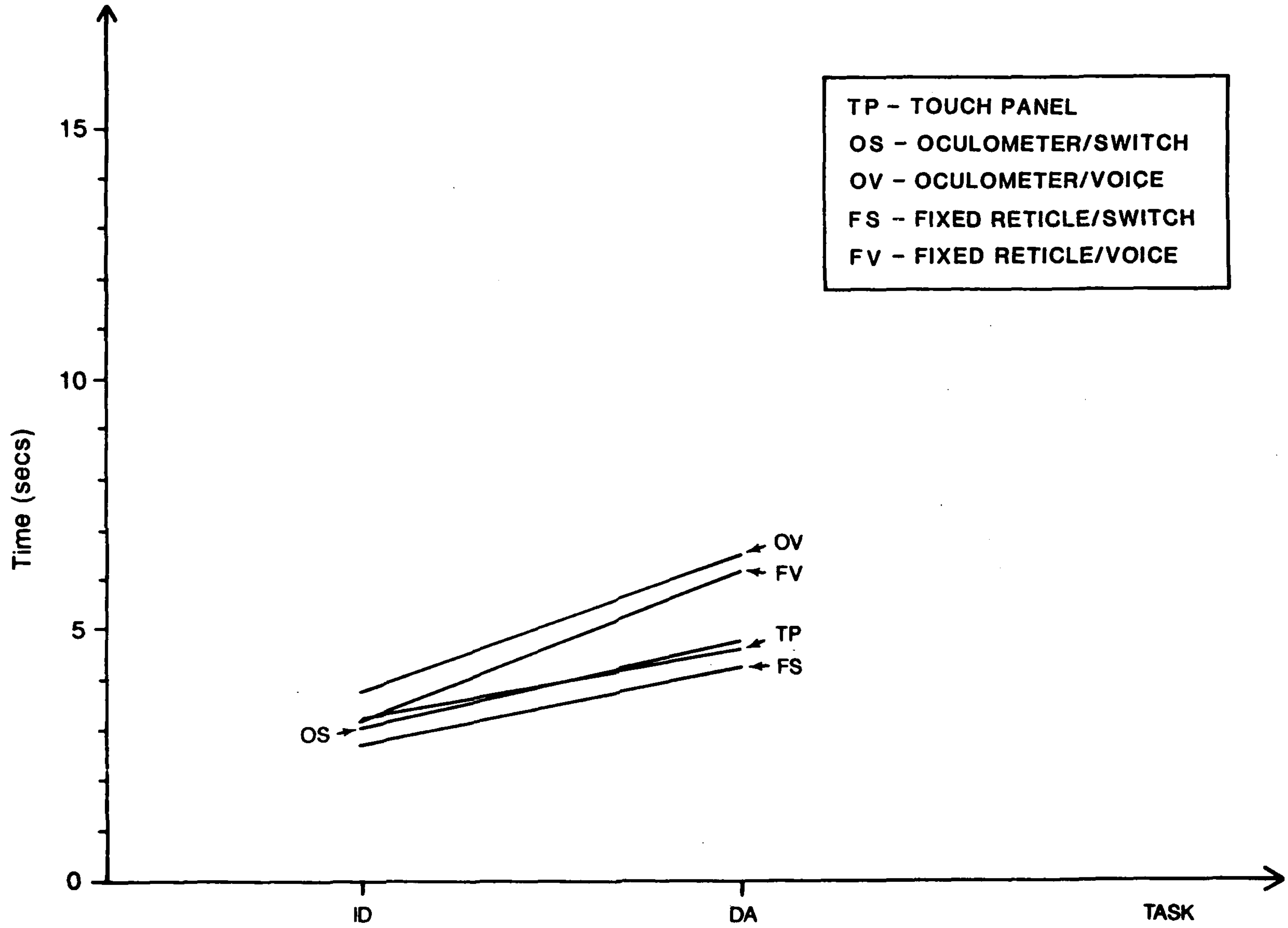
*NOTE: THE DATA VALUES FOR THIS GRAPH ARE FROM TABLE 8.

Figure 16. Task by method interaction for track select task time (T1).



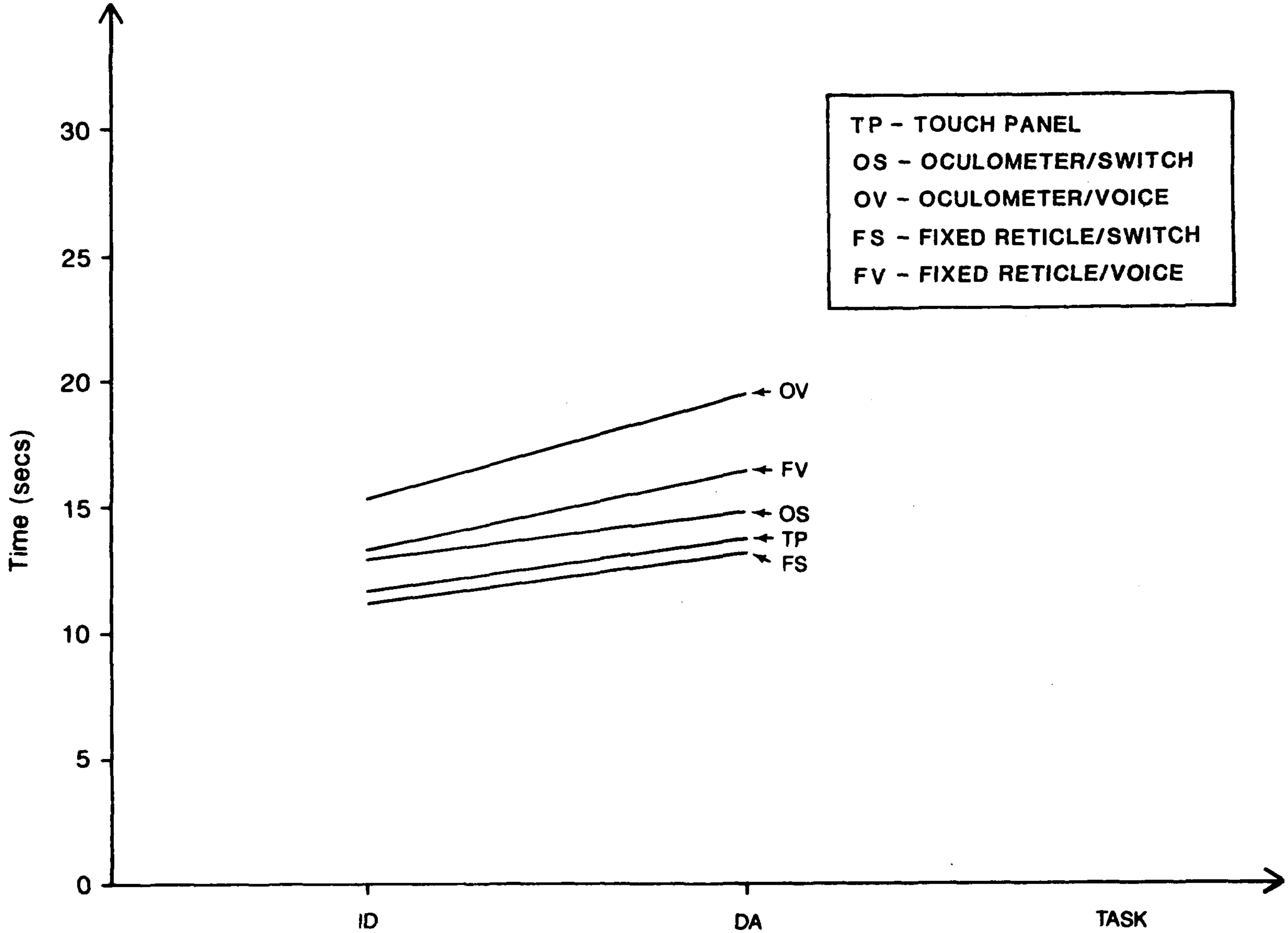
*NOTE: THE DATA VALUES FOR THIS GRAPH ARE FROM TABLE 8.

Figure 17. Task by method interaction for submenu select task time (T2).



*NOTE: THE DATA VALUES FOR THIS GRAPH ARE FROM TABLE 8.

Figure 18. Task by method interaction for subtask completion time (T3).



*NOTE: THE DATA VALUES FOR THIS GRAPH ARE FROM TABLE 8.

Figure 19. Task by method interaction for total task time (TT).

Table 11
Statistical Analysis of Subject Errors by Error Type

(a) Number of subject errors by method and error type

Method	Completed Errors: Track	ID	DA	Task	Selection Errors: Hook	Cancel
TP	0	1	1	0	1	0
FV	1	0	0	0	32	10
FS	7	0	2	0	31	8
OV	8	0	1	0	71	19
OS	8	2	0	0	49	12
Total	24	3	4	0	184	49

(b) Chi-square test of interaction between methods and error type: Contingency matrices

Observed Errors

Method	Selection: Hook	Cancel	Completed	Total
FV	32	10	1	43
FS	31	8	9	48
OV	71	19	9	99
OS	49	12	10	71
Total	183	49	29	261

Expected Errors

Method	Selection: Hook	Cancel	Completed
FV	30.15	8.07	4.78
FS	33.66	9.01	5.33
OV	69.41	18.59	11.00
OS	49.78	13.33	7.89

Test value = 7.523

Chi-square statistic = 12.59

Confidence level = .05

Degrees of freedom = 6

(Continued)

Table 11 (Continued)

(c) Chi-square test of total errors by type

	Selection			
	Hook	Cancel	**Completed**	**Total**
Observed:	184	49	31	264
Expected:	87	87	87	

Test value = 161.19

Chi-square statistic = 5.99

Confidence level = .05

Degrees of freedom = 2

KEY
TP - TOUCH PANEL
FV - FIXED RETICLE WITH VOICE ENTRY
FS - FIXED RETICLE WITH SWITCH ENTRY
OV - OCULOMETER WITH VOICE ENTRY
OS - OCULOMETER WITH SWITCH ENTRY

identification report task. Four errors were made by reporting the wrong data on the data request task. No wrong tasks were completed.

A study of the selection error data listed in Table 11 suggests that the oculometer methods, and particularly the oculometer voice method, require much more activity than do the other methods to operate. In contrast, the touch panel requires virtually no excessive activity. The fixed reticle with both voice or switch perform at about a level intermediate between the other two groupings. The selection errors for the fixed reticle methods occur roughly on the average once every fourth run. Those for the oculometer switch occur once every 2-1/2 runs, while those for the oculometer voice occur once every 1-1/2 runs. In all, 88.25% of the total errors were selection errors, and of these, 78.96% were hook action errors.

A nonparametric chi-square statistical test of the interaction between methods and error types shows no significant interaction. The contingency matrices are listed in Table 11 for both the observed errors and the expected errors. The analysis is limited to the ocular methods. The touch panel errors are too few to be included in a chi-square analysis since the corresponding expected cell values are less than unity in violation of proper test conditions (Siegel, 1956). For the same reason, the different types of completed errors are collapsed to a single column.

A large number of the errors are selection errors; about 90% of the errors are of this type. A chi-square test shows significant difference (0.05 confidence level) among the total errors for the hook selections, cancel selections, and completed errors (see Table 11). Table 12, which lists the selection errors as a function of subtask, shows that more than 80% of the hook selection errors occurred in the select menu subtask. Also, more than 76% of the cancel selections occurred in the select menu subtask. The cancel selections during the select menu subtask are probably in response to erroneously hooked tracks in the track select subtask, however. The selection of a track caused the display to automatically update from the track select to the menu select. The 14% cancel selections in the track select subtask might be in response to the "error" message that followed hook actions of tracks not in the display selection window.

A nonparametric chi-square statistical test of the interaction between methods and subtasks shows significant interaction for the hook selection errors. The contingency matrices are listed in Table 12 for both the observed errors and the expected errors. The analysis is limited to the ocular methods. Again, the touch panel errors are too few to be included in a chi-square analysis since the corresponding expected cell values are less than unity in violation of proper test conditions (Siegel, 1956). For the same reason, the data for the select submenu and complete submenu subtasks are collapsed to a single column. A chi-square test shows significant difference (0.05 confidence level) among methods for the hook selection errors of the track select subtask (see Table 12). The implication is that significantly more selection errors occurred with the oculometer voice method in the select-track subtask.

Table 12
Statistical Analysis of the Selection Errors as a Function of Subtask

(a) Selection errors as a function of method and subtask

Hook Selection Errors

Method	Select track	Select submenu	Complete submenu	Complete task
TP	1	0	0	1
FV	31	1	0	32
FS	30	1	0	31
OV	57	12	2	71
OS	34	13	2	49
Total	153	27	4	184

Cancel Selections

Method	Select track	Select submenu	Complete submenu	Complete task
TP	0	0	0	0
FV	3	7	0	10
FS	0	7	1	8
OV	3	13	3	19
OS	1	10	1	12
Total	7	37	5	49

(b) Chi-square test of method by subtask interaction for hook selection errors.

Observed
Hook Selection Errors

Method	Select-track	Select-complete	Total
FV	31	1	32
FS	30	1	31
OV	57	14	71
OS	34	15	49
Total	152	31	183

Expected
Hook Selection Errors

Method	Select-track	Select-complete
FV	26.57	5.42
FS	25.75	5.25
OV	58.97	12.03
OS	40.70	8.30

Test value = 15.37
Chi-square = 7.82
Significance level = 0.05
Degrees of freedom = 3

(Continued)

Table 12 (Continued)

(c) Chi-square test of hook selection errors on track select subtask

Hook Selection Errors
Method

	FV	FS	OV	OS	Total
Observed:	31	30	57	34	152
Expected:	38	38	38	38	

Test value = 12.89
Chi-square statistic = 7.82
Significance level = 0.05
Degrees of freedom = 3

KEY

TP - TOUCH PANEL
FV - FIXED RETICLE WITH VOICE ENTRY
FS - FIXED RETICLE WITH SWITCH ENTRY
OV - OCULOMETER WITH VOICE ENTRY
OS - OCULOMETER WITH SWITCH ENTRY

Relation of Select-Track Subtask Times and Errors

Table 13 lists the number of runs for combinations of the select-track subtask times (seconds) and hook select errors for the different methods. The time is blocked in increments of 5 seconds to 30 seconds; the last column is for task times greater than 30 seconds. Table 13 shows a uniform pattern from 0 to 10 seconds for all runs with one or no errors. There is a scattering of runs for larger errors and times beyond 10 to 15 seconds, however. This is especially true of the oculometer voice method with a few runs longer than 30 seconds and with five to six errors.

The computation of box plot outer fence values, which was derived with median-based exploratory data analysis techniques (Velleman & Hoaglin, 1983), shows that depending on the method, task times longer than 10 to 15 seconds may be considered as distribution outliers. The correlation values for the conventional Pearson's linear correlation analysis of the times to errors for the methods are listed in Table 13. The strongest correlation is for the oculometer voice method; the 0.7877 value indicates a high correlation corresponding to a marked relationship. This is because of the few outliers of long task times and many errors. A review of the data shows that these outlier runs are well scattered among the subjects and not the product of one subject or test sequence.

Sources of Selection Errors

The sources of selection errors for the ocular-based methods were probably (a) a slightly small display selection window, and (b) slight shifts in boresight alignment from the calibration settings. The size of the display selection window was determined from the published accuracies for the equipment (see Methodology, Apparatus, Display Selection Window section), and confirmed in a pilot study. The listing in Table 13 shows that runs with one selection error are intermixed with nonerror runs, however. This suggests that the selection window should have been slightly larger for the test population. Test runs with two or more errors are mainly outliers possibly caused by slight shifts in boresight alignment. While these explanations are speculatively based on observations, two questions of interest are (a) what is the cause of the misalignment, and (b) what is there about the oculometer voice method that causes the statistically significant increase in selection errors presumably associated with the misalignment? In this section of the report, we review observations on the spatial tolerance needed to maintain alignment and possible causes of boresight misalignment.

The spatial tolerances required to operate the different methods are determined by the location accuracy of the devices comprising the method and the effect of the subject's behavior on that accuracy. The accuracy of the spatial locating process for the fixed reticle methods is determined by the accuracy of the Polhemus Isotrak and the accuracy with which the reticle may be centered over the target image. The accuracy of the oculometer methods is determined by the accuracies of the Polhemus Isotrak and the NAC oculometer. These accuracies and the constraints on operation are covered in the Methodology, Apparatus section, and the Training and Test Procedures section. As noted in the Training and Test Procedures section, slight changes in facial

Table 13
Statistical Analysis of the Relationship Between the Times and Errors of the Select-Track Subtask

(a) Number of runs by task time and hook select errors for select-track subtask of each method

Method	Number errors	0-5	5-10	10-15	15-20	20-25	25-30	30+
		Subtask Time (seconds)						
TP	0	68	51	1 a	0	0	0	0
	1	4	18	5	1	0	0	0
	2	0	0	2	0	0	0	0
FV	0	48	72	0 b	0	0	0	0
	1	2	20	2	4	1	0	0
	2	0	0	0	0	0	1	0
FS	0	74	48	2 c	0	0	0	0
	1	4	12	5	0	1	0	0
	2	0	1	3	0	0	0	0
OV	0	41	75	1 d	0	0	0	0
	1	0	14	8	1	0	0	0
	2	0	0	0	2	2	0	0
	3	0	0	1	0	0	1	0
	4	0	0	0	0	0	1	0
	5	0	0	0	0	0	0	2
	6	0	0	0	0	0	0	1
OS	0	62	56	3 e	1	0	0	0
	1	2	14	4	1	1	0	0
	2	0	0	3	2	1	0	0

Note: Outer fence values for exploratory data analysis box plots:
a- 10-.39, b- 13.59, c- 12.44, d- 15.04, e- 13.48.

(b) Correlation of number of hook selection errors and task time for the select-track subtask of each method

Method	Correlation
TP	0.3602
FV	0.3759
FS	0.3351
OV	0.7877
OS	0.4636

KEY

- TP - TOUCH PANEL
- FV - FIXED RETICLE WITH VOICE ENTRY
- FS - FIXED RETICLE WITH SWITCH ENTRY
- OV - OCULOMETER WITH VOICE ENTRY
- OS - OCULOMETER WITH SWITCH ENTRY

expressions and head movements can cause a shift in the position of the oculometer mask on the face. A 0.1-millimeter shift in mask position will cause a 1.4^{o} shift in boresight (NAC, 1987).

Another influence on spatial accuracy is the behavior of the subjects when operating the different data entry methods. Some subjects pushed the data entry switch with such vigor that the reaction along the mechanical linkage between their hands and shoulders would cause a shift in the position of their heads. A change in switch type (toggle versus pressure) and position may correct this problem.

A source of error for the oculometer voice method may be the 1/2-second delay in verbalizing the voice command and processing by the automatic speech recognizer. Shifting attention from the visual modality to the auditory during this time may allow a slight relaxing of the head position. Furthermore, the head may tend to shift as the subject expels air from the lungs when voicing the verbal command. The changes in facial muscles with speech generation may also cause slight shifts in the mask on the face. One solution may be to keep a running record of previous gaze points and extract backward to the time the utterance began, using an utterance start-detection circuit interfaced to the computer.

In contrast, the fixed reticle and voice method incorporates a visual alignment task with definite visual cues and feedback from the reticle. In this case, the subject's attention may be time-shared between the visual and auditory modalities, and the 1/2-second delay in processing the verbalization would not generate head relaxation with the associated selection errors.

As they made a button push or verbal command, some subjects tended to shift their gaze to the next part of the task before completing contact. The subjects had to be instructed to hold their gaze until display-generated feedback showed that the selected action had been implemented. A faster sampling rate and maintaining a running record of previous gaze positions coupled with a backward ~action may afford a quicker and more flexible system. It is doubtful that the gaze and decision times are constants, and the implementation of such a scheme may generate erroneous selections.

Speech Recognizer Errors

The performance of the automatic speech recognizer was greater than 99.44%; it was not a source of error in operating the fixed reticle or oculometer with voice. In this test, only five nonrecognition errors were made with the automatic speech recognizer by 3 subjects of the 15. All nonrecognitions were made with the oculometer voice method. No incorrect representations or substitutions occurred. The performance of the speech recognizer was better than 99.44% since more than 900 voice command entries were made during the test. A check of the nonerror task times for the oculometer voice method shows no noticeable change when the times of the corresponding five test runs are adjusted for nonrecognitions.

The high accuracy of the speech recognizer was attained by the choice of only two command words (i.e., hook and cancel) with different spectral contents, and the setting of the reject threshold to 17, an extremely low value. The recognition score was typically in the 30 to 40 value range.

This level of recognizer performance has not been attained in previous experiments (Smyth, Denny, & Dotson, 1987) in which the recognizer was used alone with many command choices. The use of the recognizer with the oculometer or fixed reticle allows the work load to be divided among the different modalities of sight and voice, thereby reducing the word templates required and consequently increasing the performance of the recognizer.

Dropped Subjects

While 15 subjects completed the test, 20 subjects started this experiment. The remaining five subjects were dropped from the test for the following reasons. Two subjects had extremely erratic eye movement patterns and were not able to hold their eye gaze steady long enough to be calibrated with the oculometer. One subject was a 17-year-old enlisted soldier who said that he got little sleep at night. Essentially, his eyes were in a continual search pattern. The other subject was 25 years old. His eye movement patterns were more controlled but not enough to allow calibration.

Two subjects elected to discontinue the test before completion because of discomfort and fatigue. One of the uncontrolled factors in this test was the tightness of the head band on the oculometer head support system (see Figures 1 and 2) used to maintain a fixed alignment on the head. Furthermore, the 1.6 pounds (720 grams) weight of the Eyemark are loaded to the front of the mask, and the uneven distribution places a strain on the user's neck muscles. Finally, the visual attention demanded by the forced gaze, required to operate the oculometer as a pointing device, is known to reduce eye blinks and thereby dry the corneas of the eyes causing ocular fatigue. Most subjects learned with practice to tolerate these discomforts, however.

The fifth subject could not complete the test because of computer equipment breakdown caused by failure of the controller board on the VAX 11/780.

Nonerror Task Times

A question of interest is the effect that the error runs had on the statistical results for the task times. Table 14 lists the differences between the average task times for the error-free runs and all runs of each subtask and the total task for each method and task combination. The table lists also the nonparametric sign test for the case of two related samples. The results show that there is a significant difference (0.05 confidence level) between the average error-free and all data times for the total task times and the subtask times except for the subtask to complete the menu action.

Since the differences between the average error-free and the all data runs are statistically significant, the next question is whether there is a significant difference among the nonerror task times for the different methods, and if so, what is the corresponding ranking? In other words, is the statistically significant difference among the all data times for the methods a result of the errors alone? Table 15 lists the results of a nonparametric Friedman two-way statistical analysis of variance by ranks for related

Table 14
Statistical Analyses of the Difference Between the Average Times for Error-Free Runs and Total Runs

(a) Differences between average times (seconds) for the error-free runs and total runs for method and task

		Subtask			
Task	**Method**	**Select track**	**Select menu**	**Complete menu**	**Complete task**
ID	TP	-.0249	-.0021	-.0026	-.0297
	FS	-.7554	-.1131	-.0419	-.9109
	FV	-1.2623	-.1176	-.0371	-1.4173
	OS	-.8805	-.6785	-.1876	-1.6658
	OV	-1.4181	-.4282	-.3193	-2.1658
DA	TP	0.0074	-.0257	0.0173	-.0009
	FS	-.4500	0.0703	0.0103	-.3699
	FV	-.6015	-.0205	-.0739	-.6964
	OS	-1.3821	-.4236	-.1876	-1.9944
	OV	-3.0259	-.3097	0.0155	-3.3141

(b) Nonparametric sign test for two related samples applied to the difference data

		Subtask			
Task	**Method**	**Select track**	**Select menu**	**Complete menu**	**Complete task**
ID	TP	-	-	-	-
	FS	-	-	-	-
	FV	-	-	-	-
	OS	-	-	-	-
	OV	-	-	-	-
DA	TP	+	-	+	-
	FS	-	+	+	-
	FV	-	-	-	-
	OS	-	-	-	-
	OV	-	-	+	-
Sign changes		1	1	3	0
Probability		0.011	0.011	0.172	0.001

KEY

METHOD	TP - TOUCH PANEL
	FV - FIXED RETICLE WITH VOICE ENTRY
	FS - FIXED RETICLE WITH SWITCH ENTRY
	OV - OCULOMETER WITH VOICE ENTRY
	OS - OCULOMETER WITH SWITCH ENTRY
TASK	DA - DATA REQUEST
	ID - IDENTIFICATION CHANGE

Table 15
Statistical Analysis of the Error-Free Times

(a) Average times for the error-free runs by method and task

Task	Method	Number	Subtask: Select track	Select menu	Complete menu	Complete task
ID	TP	73	5.2359	3.0428	3.3859	11.6830
	FS	61	4.7633	2.9988	2.7052	10.4651
	FV	59	5.1261	3.5338	3.2970	11.9545
	OS	53	5.0882	3.3214	3.0039	11.4119
	OV	49	5.6271	4.0004	3.5070	13.1334
DA	TP	74	5.6857	3.4751	4.6381	13.7981
	FS	56	5.2758	3.2157	4.2958	12.7861
	FV	60	5.9179	3.8240	6.0959	15.8382
	OS	53	5.2346	3.2365	4.5152	12.9837
	OV	52	5.9412	3.9089	6.5165	16.3646

(b) Friedman two-way analysis of variance by ranks for related samples of the error-free times

Subtask	Task	Method: TP	FS	FV	OS	OV
Select track	ID	4	1	3	2	5
	DA	3	2	4	1	5
Select menu	ID	2	1	4	3	5
	DA	3	1	4	2	5
Complete menu	ID	4	1	3	2	5
	DA	3	1	4	2	5
Complete task	ID	3	1	4	2	5
	DA	3	1	4	2	5
	Total	25	9	30	16	40

Friedman test value = 29.1
Chi-square statistic = 9.49
Confidence level = .05
Degrees of freedom = 4

KEY	
METHOD	TP - TOUCH PANEL
	FV - FIXED RETICLE WITH VOICE ENTRY
	FS - FIXED RETICLE WITH SWITCH ENTRY
	OV - OCULOMETER WITH VOICE ENTRY
	OS - OCULOMETER WITH SWITCH ENTRY
TASK	DA - DATA REQUEST
	ID - IDENTIFICATION CHANGE

samples. The table lists the average task times for the error-free runs by method and subtask and total task for each task. Table 15 also lists the ranking of the times for each method. The difference among the methods remains statistically significant (0.05 confidence level); the rankings continue to agree with the parametric analysis of all the data times reported in the Results section. This result implies that the time differences are because of the different methods.

Nonerror Spatial Tolerance

Table 16 lists the descriptive statistics for the spatial tolerance required to operate the different methods in the track selection subtask for the nonerror trials. The statistics are the number of nonerror trials, the average displacement of the computed selection point from the selected display item, the minimum and maximum displacements, and the standard deviation. The selection point for the touch panel is the point where the subject touched the screen. The selection point for the remaining methods is the computed gaze point on the display. The average, minimum, and maximum are the appropriate statistics for the displacement of the subject's touchpoint or computed gaze point from the center of the selected track symbol. The standard deviation is that of the displacements about the center of the selected display item. The maximum value listed in Table 16 is used as an estimate of the tolerance needed to operate the method.

The track symbol for the track selection subtask is a small target (.228 $in.^2$) subtending 25 minutes of arc at the viewer's eye. The display is normal to the viewing direction. The viewing distance is about 27 inches. Table 16 lists the statistics for both inches of displacement on the display surface and the equivalent angular displacement in degrees. The concept of an angular accuracy for the touch panel as measured from the viewpoint is operationally meaningless but is included for comparison to the ocular-based methods.

An interpretation of the data listed in Table 16 is that for the track selection subtask, the touch panel requires half the tolerance required for the other devices. The tolerances required for the other devices are about the same value. The tolerance of the touch panel is about $\pm 1/2$ inch ($\pm 1^{\circ}$). The tolerance of the remaining devices is about ± 1.1 inches ($\pm 2.4^{\circ}$).

The computed tolerance of the ocular-based methods is about half the size of the display window and the actual required tolerance may be larger. The attempt to select track symbols outside the display window would be registered as selection errors.

Table 17 lists the frequency distribution for the spatial tolerance of each method. The table lists the midpoints for the spatial tolerance ranges of the eight cells from 0.0 inch to 1.384 inches and the corresponding frequencies for each spatial cell. Figure 20 is a frequency plot of the data listed in the table.

The frequency plot for the spatial tolerance of the touch panel is well within the limits of the display selection window (± 1.1 in.). In contrast, the distribution plots for the ocular methods are much wider. The plots for the switch methods are certainly skewed and possibly bimodal with the larger

Table 16
Descriptive Statistics for the Spatial Tolerance of the Track Selection Subtask Number, Average, Minimum (MIN), Maximum (MAX), and Standard Deviation (SD)

I. Display tolerance (inches)

Method	Number	Average	MIN	MAX	SD
TP	107	0.219	0.011	0.553	0.249
FV	117	0.502	0.014	1.115	0.554
FS	114	0.521	0.085	1.148	0.579
OV	88	0.516	0.076	1.226	0.584
OS	99	0.520	0.085	1.031	0.566

II. Angular tolerance (degrees)

Method	Number	Average	MIN	MAX	SD
TP	107	0.464	0.022	1.173	0.528
FV	117	1.066	0.029	2.365	1.175
FS	114	1.105	0.181	2.435	1.229
OV	88	1.096	0.161	2.600	1.240
OS	99	1.103	0.181	2.187	1.202

KEY

- TP - TOUCH PANEL
- FV - FIXED RETICLE WITH VOICE ENTRY
- FS - FIXED RETICLE WITH SWITCH ENTRY
- OV - OCULOMETER WITH VOICE ENTRY
- OS - OCULOMETER WITH SWITCH ENTRY

Table 17
Frequency Distribution of the Spatial Tolerance of Each Method

Method	Runs	Spatial Tolerance (inches)							
		0.081	0.254	0.427	0.601	0.773	0.946	1.119	1.292
TP	107	.411	.439	.140	.009	.000	.000	.000	.000
FV	117	.086	.179	.274	.248	.154	.043	.017	.000
FS	114	.079	.193	.272	.237	.088	.096	.035	.000
OV	88	.114	.171	.239	.204	.148	.079	.028	.028
OS	99	.061	.152	.263	.333	.101	.091	.000	.000

KEY

- TP - TOUCH PANEL
- FV - FIXED RETICLE WITH VOICE ENTRY
- FS - FIXED RETICLE WITH SWITCH ENTRY
- OV - OCULOMETER WITH VOICE ENTRY
- OS - OCULOMETER WITH SWITCH ENTRY

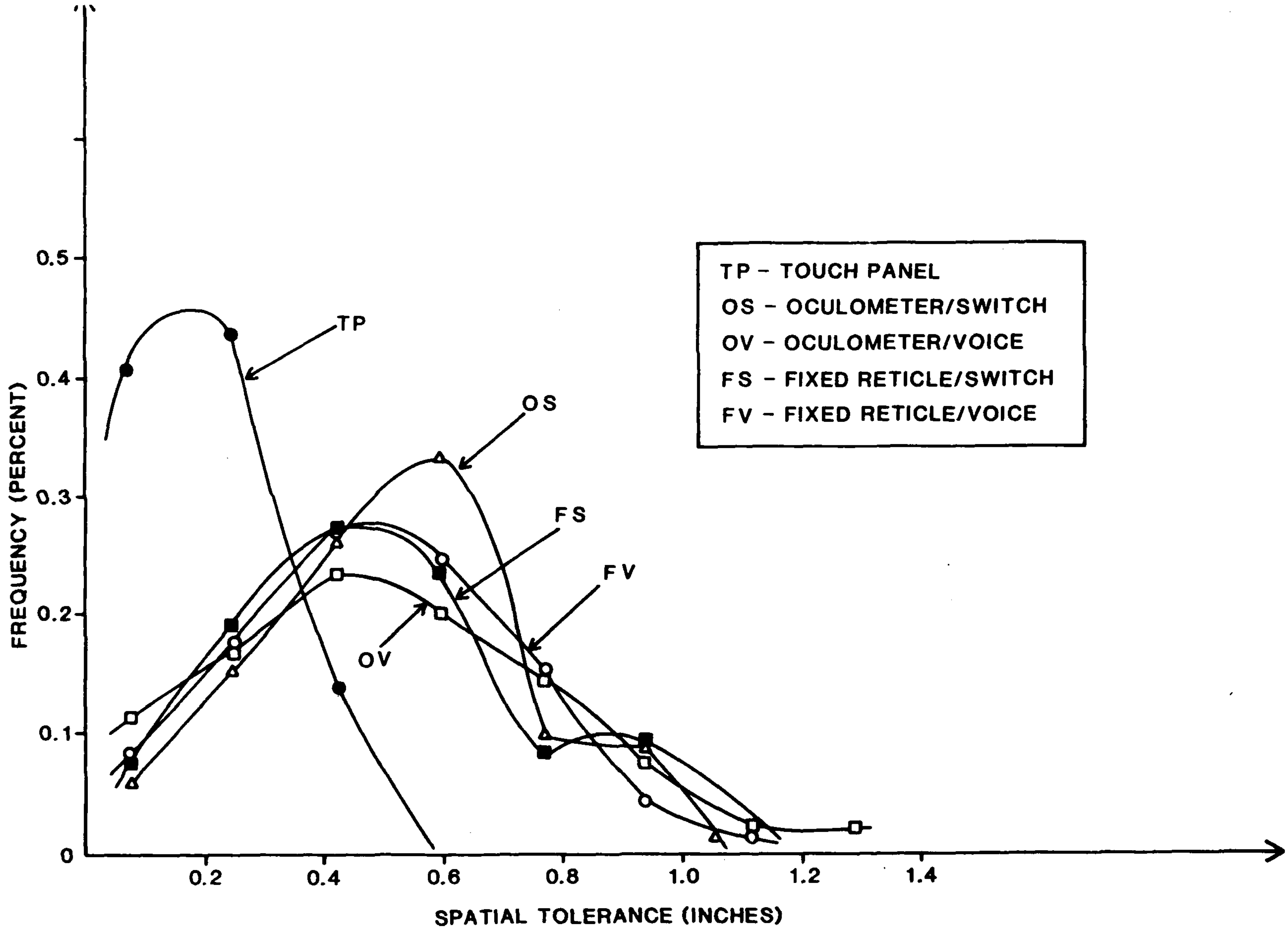


Figure 20. Spatial tolerance frequency distribution by method.

peak well within the display window and the smaller peak at the window edge. It is as though in most cases, the subjects had no difficulty in selecting tracks but would occasionally lose alignment.

The plots for the voice entry methods are more strongly skewed. They appear to be skewed unimodal with a wider distribution than the dominant peaks of the switch methods. In particular, while the plot for the reticle voice method is within the display window size, the plot for the oculometer voice appears to be slighly truncated by the display window.

Table 18 lists the descriptive statistics of the error-free spatial tolerance distributions for the different methods. The table lists the average and standard deviation from Table 16 and the peak and upper end of the distribution less than 0.05 from Table 17. No statistical analysis is made to compare the spatial distributions; the data are too few for a proper power efficiency value for the nonparametric Kolmogorov-Smirnov two-sample test of cumulative distributions.

Table 18
Descriptive Statistics of the Spatial Tolerance Distributions for the Methods

Methods	Numbers	Average	Standard deviation	Peak	Tail < .05
TP	107	0.219	0.249	0.234	0.601
FV	117	0.502	0.554	0.427	0.946
FS	114	0.521	0.579	0.427	1.119
OV	88	0.516	0.584	0.427	1.119
OS	99	0.520	0.566	0.601	1.119

KEY
TP - TOUCH PANEL
FV - FIXED RETICLE WITH VOICE ENTRY
FS - FIXED RETICLE WITH SWITCH ENTRY
OV - OCULOMETER WITH VOICE ENTRY
OS - OCULOMETER WITH SWITCH ENTRY

CONCLUSIONS

The results of this experiment suggest that the ocular-based gaze pointing methods are more inaccurate and no faster than the touch panel for selection of displayed data items in simple primary tasks involving display interactions. The touch panel and the switch-based methods are significantly faster than the voice entry methods. The head-fixed reticle methods are faster than the oculometer methods of the same modality. The gaze pointing methods require a larger display selection window than is required for the touch panel. This is especially true for the oculometer voice method which generated significantly more selection errors than the other methods, suggesting that the oculometer voice method is more inaccurate than the other methods. The ocular-based gaze pointing methods may be acceptable as a means of interacting with a display for secondary discrete tasks during flying when only a few widely spaced selections are displayed on the screen at one time.

RECOMMENDATIONS FOR FURTHER RESEARCH

The use of the ocular-based gaze pointing methods with properly designed displays may allow the display interaction part of the pilot's work load to be separated into a sequence of short discrete steps. These may cause minimal interference with the continual tracking task component used to pilot a helicopter. The pilot would interact with the displays through a few, widely spaced menu options which he would select with a momentary eye gaze and button push or voice command. Alternately, he could use a sequence of switch actions to index to the data item of interest within a display selection window centered by an initial eye gaze. For this reason, further research into the use of eye gaze techniques for display and instrument control is recommended as follows:

1. Better equipment is needed for further investigation, including a lighter, more accurate, and stable head-held oculometer and a faster computer processing system.

2. Investigate the use of responses from intrinsic variables to replace switches or voice entry as methods to select display items. The variables could include eye dwell time, pupil dilation, eye blinks, and Electroencephalograms (EEG). The processing of these signals may allow a more natural and quicker response freeing the manual modality in helicopters for piloting.

3. Investigate helicopter display formats appropriate for the reduced spatial accuracy of ocular-based gaze pointing methods. The use of widely separated menu tree selections could be used with a sequence of choices. Another approach would a sequential hook mechanism using a three-way toggle switch with positions for initialization, selection, and confirmation. The first switch action would center a selection window at the gaze point and the remaining actions would index the hooking among the track symbols within the window until a confirmation switch action is made.

4. Investigate the change in work load caused by the increase in visual loading associated with the use of such devices. The necessity of directed gazes decreases eye blinks while increasing ocular fatigue.

5. Investigate the integration of an ocular-based gaze pointing method with helmet-mounted displays (HMD). In particular, can a fixed reticle displayed on a HMD be used to interact with panel-mounted displays? Alternately, does the addition of an oculometer to a helmet allow interaction with HMDs? How can the processing system be designed to distinguish between the selection of a data item on the HMD ocular display and that of an external target when the pilot is operating in a heads-up mode?

6. Investigate the effects of the operational environment with natural sunlight and the vibrations induced in a moving helicopter on the performance of an oculometer-based system.

REFERENCES

Billingsley, P. A. (1982). Navigation through hierarchical menu structures: Does it help to have a map? Proceedings of the Human Factors Society 26th Annual Meeting, 103-107.

Breglia, D. R. (1981). Helmet Mounted Eye Tracker Using a Position Sensing Detector, PAT-APPL-6-262 153 (Patent Application), Washington, DC: Department of the Navy.

Department of Defense. (1981, May). Military Standard on Human Engineering Design Criteria for Military Systems, Equipment and Facilities (MIL-STD-1472C). Washington, DC: U.S. Government Printing Office.

Department of Defense. (1983, May). Military Standard on Symbols for Army Air System Displays (DoD-STD-1477 [MI]). Washington, DC: U.S. Government Printing Office.

Fallesen, J. J., Smyth, C. C., & Blackmer, R. F. (1983). Human Engineering Laboratory Division Air Defense Systems-I (HELDADS-I): Baseline Air Battle Management Operations Center (ABMOC) Manual Performance (Technical Memorandum 11-83). Aberdeen Proving Ground, MD: U.S. Human Engineering Laboratory.

List, Uwe H. (1983). Nonlinear prediction of head movements for helmet-mounted displays (AFHRL-TP-83-45). Williams Air Force Base, AZ: Operations Training Division, Air Force Human Resources Laboratory.

Meyer, A. U. (1981). Dynamics of two-dimensional eye-head tracking (AFOSR-TR-81-0610). Newark, NJ: Department of Electrical Engineering, Institute of Technology.

Miller, D. P. (1981). The depth/breadth tradeoff in hierarchical computer menus. Proceedings of the Human Factors Society 25th Annual Meeting, 542-546.

NAC, Inc. (1987). Operational manual: Eye Mark Recorder (Model V). NAC, Incorporated.

Setterholm, J. M. (1982). Assessment of stereo-graphics for fire control and navigation in fighter aircraft (AFWAL-TR-82-3008). Minneapolis, MN: Honeywell, Incorporated.

Siegel, S. (1956). Non-parametric statistics for the behavior sciences, New York, NY: McGraw-Hill Book Company.

Smyth, C. C., Denny, S. M., & Dotson, D. D. (1987). Comparison of voice recognition, touch panel, and keypad techniques of data entry for a forward area air defense command, control, and intelligence (FAAD C^2I) display (Technical Memorandum 25-87). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.

APPENDIX

POSTTEST QUESTIONNAIRE AND RESULTS

POSTTEST QUESTIONNAIRE

1. How would you rank the five hooking test procedures?

   The combinations were

   a. Fixed Reticle Search with Switch Input
   b. Fixed Reticle Search with Voice Input
   c. Oculometer Search with Switch Input
   d. Oculometer Search with Voice Input
   e. Touch Panel

Rank	Combination
1	
2	
3	
4	
5	

2. Why did you like 1 the best?

3. Why did you like 5 the least?

4. Do you have any suggestions for improvements?

   Overall-

   Oculometer related-

   Fixed Reticle related-

   Data Input Methods-

      Switch-

      Voice-

   Touch Panel-

5. What is your overall opinion of the experiment?

6. Do you have any suggestions for further research?

7. What is your opinion of the concept of the fixed reticle, the oculometer, and the touch panel regardless of the equipment used?

8. What improvements (in the equipment or procedure) do you feel would be required in order to make the concept of these methods more attractive?

QUESTIONNAIRE RESULTS

General problems and suggestions for improvements were noted in the subjects' questionnaires, as follow:

1. HELMET

Suggestions:

- Need a better fitting helmet to reduce fatigue and pain
- Lighten helmet
- Need better helmet adjustability
- Make mask part bigger - eliminate closed in feeling
- Need to work on better head gear

2. OCULOMETER

Suggestions:

- Should be able to move your head, too fatiguing to hold it still
- Use a head brace to relieve muscle fatigue
- Oculometer should be calibrated by the operator
- Should have a better balanced head set
- Get rid of helmet and find another method or make helmet comfortable

3. FIXED RETICLE

Problems:

- Reticle was too difficult to see
- Had trouble adjusting eye focus on target and reticle overlay
- Had trouble seeing reticle with both eyes open

Suggestions:

- Reticle should be projected to infinity and sharper image
- Reticle should be thinner
- Illuminate the reticle
- Replace cross hairs with a dot
- Use a larger reticle and adjust it to individual's focal point

4. SWITCH

Suggestions:

- Make switches more sensitive
- Get better switches
- Use hand-held switches

5. VOICE INTERACTION

Suggestions:

- Improve voice recognition given stress
- Improve voice recognition given voice inflection

6. TOUCH PANEL

Problem:

- Too cumbersome and slow

Suggestions:

- Use a soft touch switch panel
- Improve touch panel so pencil need not be put straight in and out
- Use finger to touch

Table 1
Cross Tabulation Table for Subjects' Rating of Preference

Configurations

Ratings (1 = Most preferred and 5 = Least preferred)	TP	FV	FS	OV	OS
1	S1, S3, S5, S9, S12 & S14	S2, S6, S8 & S10	S7, S11 & S15	S4 & S13	None
2	S2	S4, S11, S13, S14 & S15	S3, S6, S8, S9 & S10	S7 & S12	S1 & S5
3	S6, S11 & S15	S1, S7, S9 & S12	S2 & S14	S5, S8 & S10	S3, S4 & S13
4	None	None	S1, S4, S5, S12 & S13	S3, S6 & S14	S2, S7, S8, S9, S10, S11 & S15
5	S4, S7, S8, S10 & S13	S3 & S5	None	S1, S2, S9, S11 & S15	S6, S12 & S14

KEY

TP - TOUCH PANEL
FV - FIXED RETICLE WITH VOICE ENTRY
FS - FIXED RETICLE WITH SWITCH ENTRY
OV - OCULOMETER WITH VOICE ENTRY
OS - OCULOMETER WITH SWITCH ENTRY